Introduction

Les statistiques est un ensemble de résultats obtenus à la suite de nombreuses observations d'un même phénomène présenté le plus souvent sous forme de tableaux ou de graphiques. Les statistiques économiques concernant la production marocaine de phosphates, la mesure des précipitations sur certaines régions pendant une année, etc, sont des exemples.

Le statisticien doit interpréter ces données, établir des relations entre les faits observés en mettant en relief celles dues à l'ordre normal des choses et en écartant celles résultant de phénomènes aberrants. Pour cela, il doit utiliser des *méthodes mathématiques* constituant ce qu'on appelle *la statistique*.

La statistique est donc un ensemble de moyens permettant d'interpréter les statistiques. Cette science ne s'applique qu'aux ensembles suffisamment nombreux. En effet, certaines régularités (lois régissant certains phénomènes) n'apparaissent qu'aux niveau d'ensembles nombreux. Les erreurs faites lors d'une expérience sur un petit nombre d'essais ont beaucoup plus d'incidences (par exemple en travaux pratiques de Physique ou Chimie).

Le champ d'application de la statistique est très vaste. Voici quelques domaines où elle joue un rôle fondamental : Physique (Physique statistique), Chimie (polymère ...), Biologie (régénération de cellules), Astronomie, Agronomie, Industrie, Économie, Démographie, ...

En général, un cours de probabilité et statistique comprend trois parties :

- Statistique descriptive,
- Calcul des probabilités,
- Statistique mathématique.

La statistique descriptive consiste à classer des données, les organiser et les présenter de façon claire, décrire ou analyser une population donnée sans tirer de conclusions pour une population plus grande.

Le calcul des probabilités constitue l'outil mathématique.

En fin, la statistique mathématique permet grace aux méthodes mathématiques de tirer des conclusions sur toute la population à partir d'un échantillon (principe des sondages).

Le cours de probabilité du module M312 traitera essentiellement du calcul des probabilités, son but est de vous familiariser avec le raisonnement probabiliste.

A ses débuts, la théorie du “calcul des probabilités” a concerné principalement l'étude et la modélisation des jeux de hasard. Les premiers travaux sont attribuées à Pascal et à Fermat (1654) sur des problèmes posés par le Chevalier de Méré, joueur professionnel et mathématicien amateur. La probabilité est définie comme le nombre de cas favorables sur le nombre de cas possibles et la solution fait souvent appel au dénombrement.

Imaginons maintenant que je joue une partie de tennis contre le Numéro 1 mondial. Si on veut bien admettre, par indulgence, qu'il y'a deux solutions pour le vainqueur (lui ou moi !!), il ne paraît pas du tout raisonnable de leur attribuer la même probabilité. Il faut donc généraliser la notion de probabilité, c'est ce qui est fait au Chapitre 1 (espace probabilisé). On y définit dans sa plus grande généralité la notion de probabilité sur un ensemble fini et même dénombrable. C'est la partie “probabilités discrètes”.

Dans certaines situations, le cadre théorique précédent est insuffisant, c'est le cas en particulier quand on s'intéresse à une mesure physique (poids, tension électrique, etc) qui prend ses valeurs dans $\mathbb{R}$ qui n'est pas dénombrable. Ce sont alors d'autres techniques qui sont employées au chapitre 7 (variables aléatoires continues) avec la notion de densité de probabilité. C'est la partie “probabilités

continues".

Chapitre 1

Espaces probabilisés

Introduction

Le calcul des probabilités est la science qui modélise "l'aléatoire", c'est à dire qui propose des modèles pour comprendre les phénomènes dus au hasard. Son but est de fournir le cadre théorique général pour déterminer des lois de répartition abstraites pouvant s'adapter plus ou moins parfaitement aux données empiriques. La modélisation du calcul des probabilités a été inventée par A. N. Kolmogorov dans un livre paru en 1933. Elle est faite à partir d'objets que nous allons décrire dans ce chapitre.

1.1 Notion d'expérience aléatoire

Il y a dans la nature des phénomènes qui se déroulent de façon déterministe (sur terre une pierre tombe toujours si elle est lâchée) et d'autres de manière aléatoire : si on lance une pièce de monnaie, on ne sait pas à l'avance si elle tombe sur Pile ou Face.

Une *expérience aléatoire* est définie comme une expérience pour laquelle les conditions ne déterminent pas un résultat unique, mais laissent place à plusieurs résultats. De plus, cette expérience peut être répétée dans les mêmes conditions autant de fois que l'on veut (on dit aussi "épreuve"). Par exemple, le comptage du nombre de particules émises par une substance radioactive dans un intervalle de temps donné, est une expérience aléatoire.

1.2 Notion de résultat

Soit $\mathcal{E}$ une expérience aléatoire. On doit, en premier lieu, recenser l'ensemble des résultats possibles de $\mathcal{E}$. Cet ensemble sera noté Ω (appelé aussi ensemble fondamental).

Exemple 1.2.1 Si l'expérience est le jet de dé, l'ensemble fondamental est donné par $\Omega = \{1, 2, 3, 4, 5, 6\}$

Exemple 1.2.2 L'univers de l'expérience aléatoire qui consiste à jeter deux dés de couleurs différentes et noter les faces qui apparaissent est l'ensemble des couples (i, j) où i est le numéro de la face supérieure du premier dé et j est celui de la face supérieure du deuxième dé : $\Omega = \{(i, j), 1 \leq i \leq 6, 1 \leq j \leq 6\}$

Le résultat d'une expérience aléatoire est un élément déterminé de Ω.

1.3 Notion d'événement

Soit $\mathcal{E}$ une expérience aléatoire et Ω l'ensemble des résultats possibles. On s'interesse au fait qu'un certain événement se trouve réalisé à la suite de l'expérience.

Un événement A sera caractérisé par le sous-ensemble A de Ω des résultats $\omega \in \Omega$ qui le réalisent. $\{\omega\}$ est appelé événement élémentaire.

Exemple 1.3.1 On lance deux dés de couleurs différentes : $\Omega = [\![1,6]\!] \times [\![1,6]\!]$. Si on considère l'événement A : « la somme des points est supérieure ou égale à dix », celui-ci est représenté par la partie $\{(6,4),(5,5),(4,6),(5,6),(6,5),(6,6)\}$.

Un événement se réalise si le résultat de l'épreuve est un des événements élémentaires qui le composent.

On établit ainsi une correspondance entre le langage des événements et le langage ensembliste. On définit sur les événements des opérations logiques correspondant aux opérations définies sur les parties d'un ensemble. Le tableau suivant montre la correspondance de vocabulaire et de notations.

Événements	Sous-ensembles	Notations
Événements A, B, C	Parties A, B, C de Ω	$A \subset \Omega$
Événement élémentaire $\{\omega\}$	Élément $\omega \in \Omega$	$\omega \in \Omega$
ω réalise l'événement A	ω est un élément de A	$\omega \in A$
Événement certain	Tous les résultats possibles	Ω
Événement impossible	Partie vide de Ω	$\emptyset \in \mathcal{P}(\Omega)$
Événement contraire	Complémentaire de A dans Ω	$\bar{A}$
A ou B	Réunion de A et B	$A \cup B$
A et B	Conjonction de A et B	$A \cap B$
A et B incompatibles	Parties disjointes A et B	$A \cap B = \emptyset$
A implique B	A contenu dans B	$A \subset B$

Exemple 1.3.2 Soit l'expérience qui consiste à jeter un dé et noter sa face supérieur et soient les événements A : « point inférieur ou égal à 3 », B : « point supérieur ou égale à 3 », P : « point pair » et I : « point impair ». Les événements A et B se réalisent simultanément si et seulement si l'événement $A \cap B = \{3\}$ se réalise. l'événement « A ou B se réalisent », ou, autrement dit, « au moins un des deux événements A et B se réalise » est déterminé par la réunion $A \cup B = \Omega$ qui est l'événement certain. Les événements P et I sont incompatibles car $P \cap I = \emptyset$. En fin, l'événement « marquer un 1 ou un 3 » implique l'événement A : $\{1,3\} \subset A$.

1.4 Définition d'un espace probabilisé

A chaque événement on associe un nombre positif compris entre 0 et 1, sa probabilité. La théorie moderne du calcul des probabilités repose sur l'axiomatique suivante

Définition 1.4.1 Soit Ω l'ensemble des résultats d'une expérience aléatoire et $\mathcal{T}$ une famille de parties de Ω (i.e, $\mathcal{T} \subset \mathcal{P}(\Omega)$). On appelle probabilité sur $(\Omega, \mathcal{T})$ une application $\mathbb{P}$ de $\mathcal{T}$ dans $[0,1]$ telle que

- $\mathbb{P}(\Omega) = 1$ (probabilité de l'événement certain est égale à 1)

- Pour tout ensemble dénombrable d'événements deux à deux incompatibles $A_1, A_2, \dots A_n \dots$ on a
$$\mathbb{P}(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_n) + \cdots$$

Le triplet $(\Omega, \mathcal{T}, \mathbb{P})$ s'appelle espace probabilisé.

Remarque 1.4.2 Dans la définition 1.4.1, la famille de parties $\mathcal{T}$ n'est pas arbitraire. Elle doit vérifier les propriétés suivantes
- $\Omega \in \mathcal{T}$
- $\forall A \in \mathcal{T}, \bar{A} \in \mathcal{T}$ (stabilité par complémentation)
- Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{T}$, $A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots$ est encore un élément de $\mathcal{T}$ (stabilité par union dénombrable).

La famille $\mathcal{T}$ choisie ainsi s'appelle *tribu* d'événements. Dans ce cours, pour un ensemble Ω discret, on prendra toujours $\mathcal{T} = \mathcal{P}(\Omega)$, ensemble de toutes les parties de Ω.

Propriétés d'une probabilité

Proposition 1.4.3 *Soit $\mathbb{P}$ une probabilité sur* $(\Omega, \mathcal{T})$

1. $\mathbb{P}(\emptyset) = 0$
2. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
3. $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$
4. $A \subset B \Rightarrow \mathbb{P}(B - A) = \mathbb{P}(B) - \mathbb{P}(A)$ et $\mathbb{P}(A) \leq \mathbb{P}(B)$,
$B - A$ *étant le complémentaire de* A dans B.

Pour la démonstration de ces propriétés, voir Feuille d'exercices n°1.

1.5 Définition pratique d'une probabilité pour Ω fini ou dénombrable

Si l'ensemble des résultats d'une expérience aléatoire est fini ou infini dénombrable (discret), la caractérisation d'une probabilité se simplifie et en particulier si CardΩ est fini, la définition 1.4.1 est équivalente à la vision empirique que l'on se fait de la notion de probabilité d'un événement en tant que "fréquence de sa réalisation".

Cas d'un univers Ω fini

Proposition 1.5.1 *Soit $\Omega = \{x_1, x_2, \dots, x_n\}$ un ensemble à n éléments. Alors toute probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ est entièrement déterminée par les valeurs $p_i = \mathbb{P}(\{x_i\})$, $1 \leq i \leq n$ (i.e, probabilités des événements élémentaires) vérifiant*

$$p_i \geq 0 \quad et \quad p_1 + p_2 + \cdots + p_n = 1$$

Démonstration. Soit $\mathbb{P}$ une probabilité sur $(\Omega, \mathcal{P}(\Omega))$, alors $p_i = \mathbb{P}(\{x_i\}) \geq 0$ et puisque les événements $\{x_1\}, \{x_2\}, \dots, \{x_n\}$ sont deux à deux incompatibles et $\{x_1\} \cup \{x_2\} \cup \cdots \cup \{x_n\} = \Omega$, on a $\mathbb{P}(\{x_1\}) + \mathbb{P}(\{x_2\}) + \cdots + \mathbb{P}(\{x_n\}) = \mathbb{P}(\Omega) = 1$, i.e, $p_1 + p_2 + \cdots + p_n = 1$. Réciproquement, soient $p_1, p_2 \dots, p_n$ des réels tels que $0 \leq p_i \leq 1$ et $\sum_{i=1}^n p_i = 1$. Montrons que les nombres p_i déterminent une probabilité. Posons $\mathbb{P}(\emptyset) = 0$ et pour $A \in \mathcal{P}(\Omega)$, $A \neq \emptyset$, $A = \{x_{i_1}, x_{i_2}, \dots, x_{i_r}\}$ (ensemble à r éléments), considérons l'application qui à chaque $A \in \mathcal{P}(\Omega)$ fait correspondre $\mathbb{P}(A) = \sum_{k=1}^r p_{i_k}$.

Il est évident que $0 \leq \mathbb{P}(A) \leq 1$. Il faut maintenant montrer que l'application ainsi définie vérifie les deux conditions de la définition 1.4.1. Premièrement, puisque $\Omega = \{x_1, x_2, \dots x_n\}$ on a $\mathbb{P}(\Omega) = \sum_{i=1}^{n} p_i = 1$. Pour vérifier la deuxième condition, il suffit de la montrer pour 2 parties de Ω, le cas d'un nombre quelconque d'événements en est une conséquence pour un univers Ω fini. Soient donc $A = \{x_{i_1}, x_{i_2}, \dots, x_{i_r}\}$ et $B = \{x_{j_1}, x_{j_2}, \dots, x_{j_s}\}$ deux parties disjointes de Ω, on a donc $A \cup B = \{x_{i_1}, x_{i_2}, \dots, x_{i_r}, x_{j_1}, x_{j_2}, \dots, x_{j_s}\}$ et $\mathbb{P}(A \cup B) = \sum_{k=1}^{r} p_{i_k} + \sum_{k=1}^{s} p_{j_k} = \mathbb{P}(A) + \mathbb{P}(B)$. $\mathbb{P}$ est bien une probabilité. □

On en conclut que pour Ω fini, il suffit de se donner $p_i = \mathbb{P}(\{x_i\})$.

Cas particulier de l'équiprobabilité. C'est le cas où pour $\Omega = \{x_1, x_2, \dots, x_n\}$, tous les événements élémentaires ont la même image par l'application $\mathbb{P}$ (i.e, la même probabilité) :

$$\mathbb{P}(\{x_i\}) = \frac{1}{n} = \frac{1}{\text{Card}\Omega}$$

On en déduit

$$\forall A \in \mathbb{P}(\Omega), \quad \mathbb{P}(A) = \frac{\text{Card}A}{\text{Card}\Omega}$$

En effet, soit $A = \{x_{i_1}, x_{i_2}, \dots, x_{i_r}\}$ un sous-ensemble de Ω de cardinal r. Puisque A est la réunion de ses singletons : $A = \{x_{i_1}\} \cup \{x_{i_2}\} \cup \dots \cup \{x_{i_r}\}$ et $\mathbb{P}$ est une probabilité, alors

$$\mathbb{P}(A) = \sum_{k=1}^{r} \mathbb{P}(\{x_{i_k}\}) = \frac{r}{n} = \frac{\text{Card}A}{\text{Card}\Omega}$$

$\mathbb{P}$ s'appelle une équiprobabilité. □
On obtient l'assertion célèbre

$$\mathbb{P}(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Dans ce cas le calcul des probabilités est alors essentiellement un problème de dénombrement.

Exemple 1.5.1 *On jette deux dés de couleurs différentes et on note leurs faces supérieures. On a* $\Omega = \{(i,j), 1 \leq i \leq 6, 1 \leq j \leq 6\}$, *$i$ est le résultat du premier dé et j celui du deuxième. Donc* $\text{Card}\Omega = 36$. *On munit cet ensemble de l'équiprobabilité. Cela sous-entend que toutes les faces d'un dé ont la même chance d'apparaître (dé bien équilibré). Soit A_k l'événement « la somme des deux dés est égale à k ». Calculer la probabilité de l'événement A_k pour chaque valeur possible de k.*

Les événements A_k se traduisent dans le langage ensembliste par les parties de Ω : $A_k = \{(i,j) \in \Omega, i+j = k\}$ pour $k = 2, 3, \dots, 12$. Par exemple $A_5 = \{(1,4), (2,3), (3,2), (4,1)\}$ et $A_9 = \{(3,6), (4,5), (5,4), (6$ sont deux parties de cardinal 4, donc $\mathbb{P}(A_5) = \mathbb{P}(A_9) = 4/36 = 1/9$. En général

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(A_k)$	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

Cas d'un univers Ω infini dénombrable

Proposition 1.5.2 *Soit $\Omega, \mathcal{T}$ un espace probabilisable avec $\Omega = \{x_i, i \in \mathbb{N}^*\}$ un ensemble infini dénombrable ($x_i \neq x_j$, pour $i \neq j$). Alors il existe une probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{P}(\Omega))$ telle que $p_i = \mathbb{P}(\{x_i\})$, si et seulement si les nombres p_i vérifient*

1. *$p_i \geq 0$ pour tout i,*
2. *La série de terme général p_i converge et sa somme est égale à 1 (i.e, $\sum_{i=1}^{\infty} p_i$ est finie et est égale à 1).*

Exemple 1.5.2 *On considère l'expérience aléatoire suivante $\mathcal{E}$: " On joue à Pile ou Face jusqu'à obtenir Pile ". Quels sont les résultats possibles de cette expérience ?*

On peut proposer le modèle suivant. Les résultats observables sont les suites formées de plusieurs Faces suivis par un Pile, symboliquement on écrit un résultat comme

$$\omega_n = (\underbrace{F, F, \ldots, F}_{n-1 \text{ fois}}, P)$$

ce qui signifie qu'on obtient Face dans les $n-1$ premiers lancers et Pile dans le $n^{\text{ième}}$. L'univers associé à cette expérience est donc

$$\Omega = \{\omega_n, n \in \mathbb{N}^*\}$$

c'est un ensemble infini dénombrable.

Soit l'application $\mathbb{P}$ qui fait correspondre à chaque événement élémentaire $\{\omega_n\}$ le nombre $\frac{1}{2^n}$. Montrer que l'on a défini une probabilité sur l'espace $(\Omega, \mathcal{P}(\Omega))$ associé à l'expérience aléatoire $\mathcal{E}$.

Il faut tout simplement montrer que les nombres $p_n = \mathbb{P}(\{\omega_n\}) = \frac{1}{2^n}$ pour $n \in \mathcal{N}^*$ vérifient les deux conditions de la proposition précédente. Premièrement, il est évident que les nombres p_n sont positifs. Deuxièmement, calculons $\sum_{n=1}^{\infty} p_n$.

$$\sum_{n=1}^{\infty} p_n = \lim_{N \to +\infty} \sum_{n=1}^{N} \frac{1}{2^n}$$

Or, la somme des N premiers termes de la suite géométrique $(p_n)_{n \in \mathbb{N}^*}$ est

$$\sum_{n=1}^{N} \frac{1}{2^n} = \frac{1}{2} \frac{1 - \left(\frac{1}{2}\right)^N}{1 - \frac{1}{2}} = 1 - \frac{1}{2^N}$$

Il s'ensuit finalement que

$$\sum_{n=1}^{\infty} p_n = \lim_{N \to +\infty} \left(1 - \frac{1}{2^N}\right) = 1$$

FIN DU CHAPITRE 1

Chapitre 2

Dénombrement

1 Principe de multiplication

C'est la règle fondamentale du dénombrement. Elle sert à dénombrer un ensemble fini en décrivant une suite d'expériences permettant de construire un élément de cet ensemble. Le principe s'énonce comme suit

Soit une suite de r expériences telles que la première expérience a n_1 résultats possibles. À chaque résultat de la première expérience il y a n_2 résultats possibles à la deuxième expérience. À chaque résultat possible de la première expérience et de la deuxième expérience il y a n_3 résultats possibles à la troisième expérience et ... ainsi de suite jusqu'à la $r^{\text{ième}}$ expérience. Alors le nombre total de résultats possibles à la fin des r expériences est $n_1 \times n_2 \times \cdots \times n_r$.

Précisons que le nombre de résultats de la $i^{\text{ième}}$ expérience doit être indépendant des expériences précédentes.

Exemple 1.1 Combien de "mots" comportant 3 lettres distinctes peut-on former ?

La première expérience qui consiste à choisir une lettre parmi 26 lettres de l'alphabet a 26 résultats ($n_1 = 26$). Pour le choix de la deuxième lettre une fois la première est choisie, il y a 25 résultats ($n_2 = 25$). Les deux premières lettres étant choisies, il reste 24 choix pour la troisième ($n_3 = 24$). La procédure indiquée permet bien de construire tous les mots à 3 lettres distinctes une fois et une fois seulement. Le principe de multiplication donne $26 \times 25 \times 24$ possibilités.

Exemple 1.2 Une association de consommateurs note un produit selon 3 critères :
- Facilité d'utilisation (F) : Bonne (F1), Moyenne (F2), Mauvaise (F3).
- Prix (P) : cher (P1), pas cher (P2).
- Coût de maintenance (C) : cher (C1), moyen (C2), pas cher (C3)

Combien y a t-il de possibilités de classement pour un produit ?

Le principe de multiplication donne $3 \times 2 \times 3 = 18$ possibilités.

Pour cet exemple, le diagramme en arbre (tree diagram) permet de visualiser le principe de multiplication.

Exemple 1.3 Dans un examen on a 12 questions Vrai/Faux, de combien de façons possibles un étudiant peut-il répondre ?

Réponse. $2 \times 2 \times \cdots \times 2 = 2^{12}$ façons.

Le diagramme en arbre serait difficile a faire dans ce cas.

2 Tirages successifs avec remise : Listes

Définition 2.1 Soit n et p deux entiers non nuls. Dans une population d'effectif n, on effectue l'expérience aléatoire qui consiste à **extraire successivement, avec remise**, p individus. Il en résulte que le même individu peut être choisi plusieurs fois et que l'on peut avoir $p \geq n$. Les résultats de ces tirages successifs, rangés dans l'ordre de leur obtention, constituent une **liste à** p **éléments** également appelée p-liste.

Exemple 2.1 Une urne contient 8 boules numérotées de 1 à 8. On en tire successivement 5, en notant après chaque tirage le numéro obtenu puis en remettant la boule tirée dans l'urne avant le tirage suivant. Le résultat obtenu, par exemple 38322, est une 5-liste. L'ordre intervient et les éléments sont **distincts ou non**.

Dénombrement. À chaque tirage, le nombre d'issues possibles est n et, puisqu'on effectue p tirages successifs, par le principe de multiplication, le nombre de listes à p éléments est $n \times n \times \cdots \times n = n^p$.

3 Tirages successifs sans remise : Arrangements

Définition 3.1 Soit n et p deux entiers non nuls. Dans une population d'effectif n, on extrait, **successivement et sans remise**, p individus. Il en résulte que le même individu ne peut pas être choisi plusieurs fois et que $p \leq n$. Ces tirages successifs sans remise sont dits tirages exhaustifs.

Les résultats, rangés dans l'ordre de leur obtention, constituent un arrangement de p éléments distincts choisis parmi n. On dit aussi arrangement de n éléments p à p.

Exemple 3.1 L'urne est celle de l'exemple 2.1, un exemple d'arrangement de 5 éléments choisis parmi 8 est 85642. L'ordre intervient et tous les éléments sont **distincts**.

Dénombrement. Le nombre d'issues possibles pour le premier tirage est n. Il n'est plus que de $n-1$ pour le deuxième, $n-2$ pour le troisième et ainsi de suite jusqu'au $p^{\text{ième}}$ tirage où le nombre de possibilités est $n-(p-1)$. Par le principe de multiplication, le nombre total d'arrangements de n éléments p à p est $n \times (n-1) \times (n-2) \times \cdots (n-p+1)$. Ce nombre est noté A_n^p :

$$A_n^p = n(n-1)(n-2)\cdots(n-p+1) = \frac{n!}{(n-p)!}$$

Exemple 3.2 On a 5 puces électroniques distincts mais interchangeables. De combien de façons peut-on les aligner ?

Réponse. $A_5^5 = 5 \times 4 \times 3 \times 2 \times 1 = 120$.

Le cas $p = n$ est un cas particulier important. C'est le nombre de façons de réarranger les n individus (boules) de la population (urne). On dit qu'il y a $A_n^n = n!$ permutations de n objets.

4 Tirage simultané : Combinaisons

Définition 4.1 Soit n et p deux entiers non nuls. Dans une population d'effectif n, on extrait **simultanément** p individus (on tire une poignée de p individus). Il en résulte que l'ordre de tirage des individus **n'est pas important** et que $p \leq n$. On parle de combinaison. Une combinaison d'un ensemble de n éléments distincts pris p à la fois est n'importe quel sous-ensemble de p éléments de cet ensemble. On dit aussi combinaison de n éléments p à p.

Exemple 4.1 Dans l'urne de l'exemple 2.1, on procède au tirage de 5 boules simultanément. Des exemples de résultat sont les poignées $\{2, 3, 5, 6, 8\}, \{1, 4, 5, 7, 8\}$. On ne tient pas compte de l'ordre, chaque résultat est un ensemble.

Dénombrement. On peut voir une combinaison p à p de n éléments comme un arrangement où l'ordre de tirage des éléments n'importe pas. Étant donné qu'il y $p!$ façons de permuter les p éléments dans un arrangement (par exemple il y a 5 ! façons d'aligner les boules dans l'arrangement 85642), le nombre de combinaisons est égale au nombre d'arrangements divisé par $p!$: $A_n^p/p!$. Ce nombre est noté par C_n^p

$$C_n^p = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}$$

La notation anglo-saxonne pour les combinaisons est un peu différente : $\binom{n}{p}$.

Exemple 4.2 De combien de façons peut-on choisir 3 assistants de laboratoire parmi 20 pour assister dans une certaine expérience ?

Réponse. C'est sans importance qui en particulier sont choisis et dans quel ordre. Il y a $C_{20}^3 = \frac{20!}{3!(20-3)!} = 1140$ façons.

Exemple 4.3 De combien de façons un directeur de laboratoire peut-il choisir 2 chimistes parmi 7 candidats et 3 physiciens parmi 9 autres candidats ?

Réponse. Le nombre de choix des chimistes est $C_7^2 = 21$ et le nombre de choix des physiciens est $C_9^3 = 84$. Par le principe de multiplication, la réponse est $21 \times 84 = 1764$.

4.1 Quelques propriétés

Parmi les très nombreuses propriétés des combinaisons, on peut noter

$$\begin{aligned}
C_n^0 &= C_n^n = 1 \\
C_n^p &= C_n^{n-p} \\
C_n^p &= C_{n-1}^{p-1} + C_{n-1}^p \quad \text{(Triangle de Pascal)} \\
(a+b)^n &= \sum_{i=0}^{n} C_n^i a^i b^{n-i} \quad \text{(Formule du binôme)}
\end{aligned}$$

En raison de la quatrième propriété, le nombre C_n^i s'appelle coefficient binomial.

4.2 Coefficients multinomiaux

Nous traiterons dans cette sous-section du problème suivant : un ensemble de n objets distincts doit être divisé en r groupes de tailles respective $n_1, n_2, \ldots, n_r$, avec $\sum_{i=1}^{r} n_i = n$. De combien de manières peut-on le faire ?

Pour le savoir, remarquons qu'il y a $C_n^{n_1}$ possibilités de choix pour le premier groupe ; pour chacun de ces choix, il y a $C_{n-n_1}^{n_2}$ possibilités de choix pour le deuxième groupe ; pour chaque choix des deux premiers groupes, il y a $C_{n-n_1-n_2}^{n_3}$ possibilités pour le troisième et ainsi de suite. En utilisant alors le principe de multiplication, il y aura

$$\begin{aligned} & C_n^{n_1} C_{n-n_1}^{n_2} C_{n-n_1-n_2}^{n_3} \cdots C_{n-n_1-n_2-\cdots-n_{r-1}}^{n_r} \\ = \; & \frac{n!}{(n-n_1)!n_1!} \frac{(n-n_1)!}{(n-n_1-n_2)!n_2!} \cdots \frac{(n-n_1-n_2-\cdots-n_{r-1})!}{0!n_r!} \\ = \; & \frac{n!}{n_1!n_2!\cdots n_r!} \end{aligned}$$

divisions possibles.

Le coefficient $\frac{n!}{n_1!n_2!\cdots n_r!}$ représente donc le nombre de répartitions possibles de n objets en r groupes distincts de taille respective $n_1, n_2, \ldots, n_r$. La notation mondialement convenue pour ces coefficients est

$$\frac{n!}{n_1!n_2!\cdots n_r!} = \binom{n}{n_1, n_2, \cdots, n_r}$$

Exemple 4.4 Le nombre de manières de répartir les 52 cartes d'un jeu de bridge en 4 mains de 13 cartes est $\frac{52!}{13!13!13!13!} = 53644737765488792839237440000$.

Exemple 4.5 Le poste de police d'une petite ville compte 10 agents. Si l'organisation de ce poste est d'avoir 5 agents en patrouille, 2 au poste travaillant activement et les 3 autres au poste également mais de réserve, à combien de répartition de ces agents en trois groupes définis peut-on procéder ?

Réponse. Il y a $\frac{10!}{5!2!3!} = 2520$ répartitions.

4.2.1 Formule du multinôme

Cette formule généralise la formule du binôme.

Théorème 4.2

$$(x_1 + x_2 + \cdots + x_r)^n = \sum \binom{n}{n_1, n_2, \cdots, n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}$$

La somme étant étendue à tous les r-uples $(n_1, n_2, \ldots, n_r)$ d'entiers naturels tels que $n_1 + n_2 + \cdots + n_r = n$.

Les coefficients $\binom{n}{n_1, n_2, \cdots, n_r}$ s'appellent **coefficients multinomiaux**.

Démonstration. La formule du multinôme peut se démontrer par récurrence (objet d'un exercice). □

Exercice. Développer $(x_1 + x_2 + x_3)^3$ et $(x_1 + x_2 + x_3 + x_4)^2$.

La formule du binôme s'obtient comme cas particulier de la formule du multinôme, pour $r = 2$; et dans ce cas les coefficients multinomiaux sont les coefficients binomiaux.

Fin du chapitre 2

Université Hassan II - Mohammadia
F.S.T Mohammadia Département de Mathématiques
Module : Probabilités et statistiques (MIP)

Chapitre 3

Conditionnement & Indépendance

1 Probabilité conditionnelle

1.1 Définition

La notion de probabilité conditionnelle s'introduit à chaque fois que pendant le déroulement d'une expérience aléatoire une ou plusieurs informations «de dernière minute» doivent être tenues en compte. L'évaluation des chances qu'un événement arrivera peut être très différente selon ces informations. Par exemple, l'évaluation de la probabilité qu'une maison s'effondrera demain devrait clairement être beaucoup plus grande, si on s'attend à un tremblement de terre violent, que cela serait s'il n'y avait aucune raison de s'attendre à une activité sismique inhabituelle.

Les deux exemples qui suivent montrent concrètement comment la probabilité d'un événement A peut changer radicalement si on se donne des informations sur la réalisation ou non d'un autre événement B.

Exemple 1.1 On jette 2 pièces de monnaie bien équilibrées. L'univers associé à cette expérience est $\Omega = \{\text{FF,FP,PF,PP}\}$. Soit A l'événement «Avoir deux faces » $= \{\text{FF}\}$ et B l'événement «Au moins une pile » $= \{\text{PP,PF,FP}\}$. Puisque tous les résultats possibles sont équiprobables $\mathbb{P}(A) = \frac{1}{4}$. Cependant si nous savons que l'événement B est réalisé, l'événement A ne peut pas se réaliser car ils sont incompatibles. Nous dirons que la probabilité de A, sachant que B est réalisé est nulle et nous écrirons $\mathbb{P}(A|B) = 0$ qui se lit « probabilité de A sachant B ».

Exemple 1.2 Le directeur d'une entreprise a les dossiers de 16 000 employés dans ses filières. Le dénombrement des dossiers en fonction de l'âge et du sexe des employés est présenté dans le tableau suivant

	Homme	Femme	Totaux
Moins de 30 ans	1200	1700	2900
30 ans à 40 ans	2600	4200	6800
Plus de 40 ans	4000	2300	6300
Totaux	7800	8200	16000

Un dossier est tiré au hasard des filières du directeur. Soit A l'événement « le dossier choisi est celui d'une femme ». Clairement, nous avons $\mathbb{P}(A) = \frac{8200}{16000}$. Soit B l'événement « le dossier choisi est celui d'une personne de 30 à 40 ans ». Si nous disons maintenant que le dossier tiré est celui d'une personne de 30 à 40 ans, alors la probabilité qu'il soit celui d'une femme est différente - cette dernière est $\mathbb{P}(A|B)$. Nous nous intéressons maintenant seulement aux dossiers des personnes âgées de 30 à 40 ans, i.e, les 6800 personnes de la deuxième ligne du tableau. Au sein de ce groupe il y a

4200 dossiers de femmes dont les choix sont équiprobables dans l'ensemble des employés âgés de 30 à 40 ans. Donc

$$\mathbb{P}(A|B) = \frac{4200}{6800} = \frac{\text{Nbre de femmes agées de 30 à 40 ans}}{\text{Nbre de personnes dans la tranche d'âge } [30, 40]}$$

En général

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Nous pouvons donc énoncer la définition suivante

Définition 1.1 Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, B un événement dont la probabilité est non nulle. Soit A un événement. On appelle probabilité de A sachant B, et on note $\mathbb{P}(A|B)$, le nombre défini par

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \tag{1}$$

Attention : « A sachant B » ou « $A|B$ » ne représente pas un événement, i.e, n'est pas une partie de Ω. C'est le piège de la notation $\mathbb{P}(A|B)$.

Proposition 1.2 L'application $\mathbb{P}_B : \mathcal{T} \to \mathbb{R}$, définie par

$$\forall A \in \mathcal{T}, \quad \mathbb{P}_B(A) = \mathbb{P}(A|B)$$

est une probabilité appelée probabilité conditionnée par l'événement B.

Démonstration. Voir T.D

Exemple 1.3 Voici un exemple classique. Considérons une famille dont nous savons qu'elle a deux enfants. Il existe donc quatre compositions possibles de la famille que nous noterons FF, FG, GF, GG (dans l'ordre de naissance) et que nous supposerons équiprobables.

• Quelle est la probabilité que les deux enfants soient des garçons sachant que l'aîné est un garçon ?

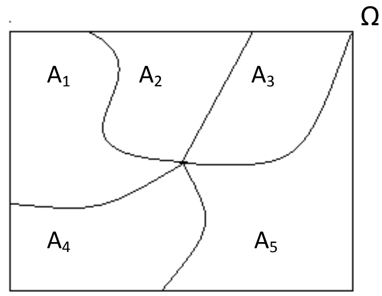
Soient
B : « l'aîné est un garçon » = {GF,GG}
A : « les deux enfants sont des garçons » = {GG}
On veut calculer $\mathbb{P}(A|B)$. Par définition

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/4}{1/2} = \frac{1}{2}$$

• Quelle est la probabilité que les deux enfants soient des garçons sachant qu'il y a au moins un garçon ?

Soit C : « il y a au moins un garçon » = {GF,FG,GG}. On cherche $\mathbb{P}(A|C)$:

$$\mathbb{P}(A|C) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{1/4}{3/4} = \frac{1}{3}$$



1.2 Formule des probabilités composées

La relation qui définit la probabilité conditionnelle peut s'écrire de la façon suivante :

$$\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}(A|B) = \mathbb{P}(A) \times \mathbb{P}(B|A) \tag{2}$$

Cette formule s'appelle *formule des probabilités composées.* Elle sert souvent à calculer la probabilité de la conjonction de deux événements en utilisant la probabilité conditionnelle.

Exemple 1.4 Une boîte renferme 3 boules blanches et 2 boules noires. On tire successivement sans remise 2 boules. Quelle est la probabilité d'obtenir la première boule noire et la deuxième noire ?

Notons A_1 : «la première boule tirée est noire» et A_2 : «la deuxième boule est noire». Nous voulons calculer $\mathbb{P}(A_1 \cap A_2)$. On a $\mathbb{P}(A_1) = \frac{2}{5}$ et $\mathbb{P}(A_2|A_1) = \frac{1}{4}$ donc

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_2|A_1)\mathbb{P}(A_1) = \frac{1}{10}$$

La formule des probabilités composées se généralise facilement (par récurrence) pour n événements. Soient $A_1, A_2, \dots A_n$ des événements tels que $\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$ alors :

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}(A_2|A_1) \times \mathbb{P}(A_3|A_1 \cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}) \tag{3}$$

Exercice. Une urne contient 4 boules blanches, 3 boules noires et 5 boules rouges. On tire une à une et sans remise 5 boules. Quelle est la probabilité que les deux premières boules tirées soient blanches, la troisième noire et les deux dernières boules soient rouges. Rép : $\frac{1}{132}$

2 Formule des probabilités totales

Commençons par donner la définition de systèmes complets d'événements.

Définition 2.1 Un système complet d'événements est une famille d'événements $(A_i)_{i \in I}$ (I étant un ensemble d'indices) qui sont incompatibles deux à deux et dont la réunion est l'ensemble Ω, c'est à dire

$$A_i \neq \emptyset, \quad \forall i \neq j, \quad A_i \cap A_j = \emptyset, \quad \text{et} \quad \bigcup_{i \in I} A_i = \Omega$$

Le dessin suivant illustre un système complet de 5 événements. On dit aussi partition de Ω.

Exemple 2.1 Soit A un événement de probabilité non nulle. le système formé de A et de son événement contraire $\overline{A}$ est un système complet.

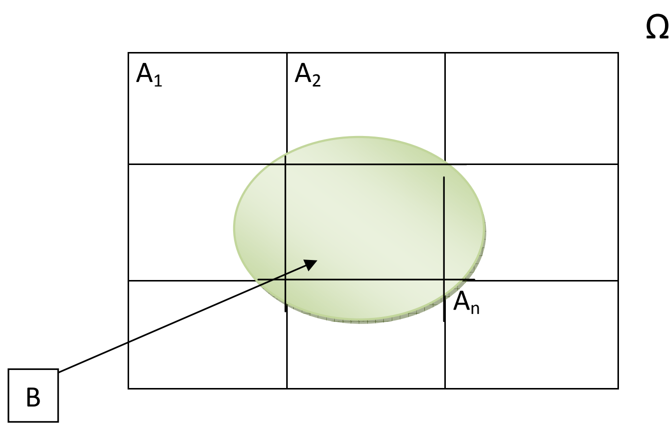
Proposition 2.2 Soient $(A_i)_{i\in I}$ un système complet d'événements d'un espace probabilisé tels que $\mathbb{P}(A_i) \neq 0, \forall i \in I$. Alors, pour tout événement B on a :

$$\mathbb{P}(B) = \mathbb{P}(A_1)\mathbb{P}(B|A_1) + \mathbb{P}(A_2)\mathbb{P}(B|A_2) + \cdots + \mathbb{P}(A_n)\mathbb{P}(B|A_n) + \cdots = \sum_{i\in I} \mathbb{P}(A_i)\mathbb{P}(B|A_i) \qquad (4)$$

Preuve. On peut écrire $\Omega = \bigcup_{i\in I} A_i$ et donc $B = \Omega \cap B = \bigcup_{i\in I}(B \cap A_i)$ car l'intersection est distributive sur la réunion (voir illustration sur la figure ci-après). Puisque les A_i sont deux à deux disjoints il en est de même pour les ensembles $B \cap A_i$. D'où

$$\mathbb{P}(B) = \sum_{i\in I} \mathbb{P}(B \cap A_i)$$

En appliquant la formule des probabilités composées (2) à chacun des termes de cette somme, on trouve la formule des probabilités totale (4). □



Remarque 2.3 La formule des probabilités totales est le plus souvent utilisée dans le cas d'un système formé d'un événement A et de son contraire : $\{A, \overline{A}\}$. On a dans ce cas

$$\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(\overline{A})\mathbb{P}(B|\overline{A})$$

Exemple 2.2 Un automobiliste arrive à proximité d'un feu tricolore. On suppose qu'aucun véhicule ne le précède. On suppose que, si le feu est vert à ce moment là, l'automobiliste décide de passer avec une probabilité de 99/100. Si le feu est orange, l'automobiliste décide de passer avec une probabilité de 3/10 et enfin si le feu est rouge, l'automobiliste décide de passer avec une probabilité de 1/100 (quelques fous...). Le cycle du feu tricolore dure une minute : vert : 25s, orange : 5s et rouge : 30s. Quelle est la probabilité que l'automobiliste passe sans s'arrêter à ce feu tricolore ?

Soient A l'événement « l'automobiliste passe sans s'arrêter au feu » et V (respectivement O et R) : « le feu est vert (respectivement orange et rouge) ». Comme le système $\{V, O, R\}$ est complet, on a

$$\mathbb{P}(A) = \mathbb{P}(A|V)\mathbb{P}(V) + \mathbb{P}(A|O)\mathbb{P}(O) + \mathbb{P}(A|R)\mathbb{P}(R) = \frac{99}{100}\frac{25}{60} + \frac{3}{10}\frac{5}{60} + \frac{1}{100}\frac{30}{60} = \frac{177}{400} < \frac{1}{2}$$

Remarque 2.4 Il est recommandé de repérer dans tout exercice de probabilité les systèmes complets qui interviennent et de leur appliquer la formule des probabilités totales.

3 Formule de Bayes

C'est une corollaire très simple et très célèbre de la formule des probabilités totales. Reprenons la formule (2)

$$\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}(A|B) = \mathbb{P}(A) \times \mathbb{P}(B|A)$$

Donc, on peut écrire pour tous événements A et B

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(B)}$$

En particulier si A est l'un des membres d'un système complet : $A = A_j$ pour un certain indice $j \in I$, on trouve

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(A_j)\mathbb{P}(B|A_j)}{\sum_{i\in I}\mathbb{P}(A_i)\mathbb{P}(B|A_i)}$$

Où nous avons appliqué la formule des probabilités totales (4) à l'événement B.

Exemple 3.1 On dispose de deux urnes U_1 et U_2. L'urne U_1 contient r_1 boules rouges et n_1 boules noires, l'urne U_2 contient r_2 boules rouges et n_2 boules noires. On tire au hasard une des deux urnes. Dans l'urne choisie on tire une boule. Sachant qu'on a tiré une boule rouge, quelles chances a-t-on de l'avoir tirée de U_1 ?

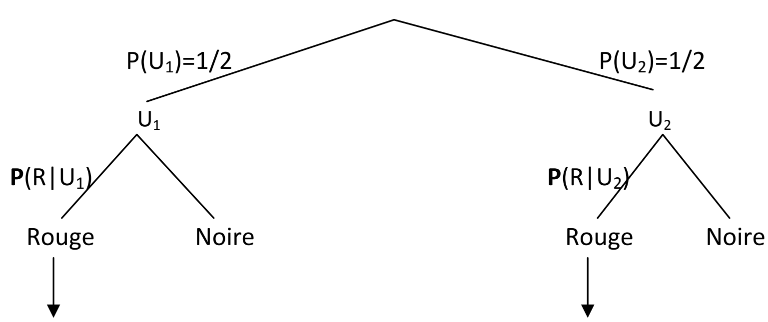
Soient les événements R : « tirer une boule rouge » et U_i : « tirer l'urne U_i » $(i = 1, 2)$. On cherche à évaluer $\mathbb{P}(U_1|R)$. Pour cela, nous allons appliquer la formule de Bayes sur le système complet $\{U_1, U_2\}$ $(U_2 = \overline{U_1})$:

$$\mathbb{P}(U_1|R) = \frac{\mathbb{P}(U_1)\mathbb{P}(R|U_1)}{\mathbb{P}(R)} = \frac{\mathbb{P}(U_1)\mathbb{P}(R|U_1)}{\mathbb{P}(U_1)\mathbb{P}(R|U_1) + \mathbb{P}(U_2)\mathbb{P}(R|U_2)}$$

Or $\mathbb{P}(U_1) = \frac{1}{2}$, $\mathbb{P}(U_2) = \frac{1}{2}$, $\mathbb{P}(R|U_1) = \frac{r_1}{r_1 + n_1}$, $\mathbb{P}(R|U_2) = \frac{r_2}{r_2 + n_2}$

Il s'ensuit en remplaçant

$$\mathbb{P}(U_1|R) = \frac{\frac{r_1}{r_1+n_1}}{\frac{r_1}{r_1+n_1} + \frac{r_2}{r_2+n_2}}$$



chch**P**(R|U$_1$) P(U$_1$) + **P**(R|U$_2$) P(U$_2$) = **P**(R)

Remarque 3.1 La formule de Bayes est indiquée lorsqu'on a une expérience aléatoire à deux niveaux. Dans l'exemple précédent, le hasard intervient dans le choix de l'urne (premier niveau) et dans le

tirage d'une boule (deuxième niveau). Un événement lié au premier niveau est la cause (choix de l'urne U_1 par exemple) tandis qu'un événement lié au deuxième niveau est la conséquence (tirage d'une boule rouge). Dans ce cas, la formule de Bayes calcule la probabilité de la cause sachant la conséquence, qui exige une remontée dans le temps, en fonction de la probabilité de la conséquence sachant la cause, qui est plus accessible.

4 Indépendance d'événements

Il est naturel de dire que deux événements A et B d'un même espace probabilisé sont indépendants si la probabilité de A sachant B réalisé, reste égale à la probabilité de A (aucune influence de la réalisation de B sur celle de A) autrement dit $\mathbb{P}(A|B) = \mathbb{P}(A)$. La formule (2) devient

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

D'où la

Définition 4.1 Deux événements d'un même espace probabilisé sont indépendants si l'on a

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B) \tag{5}$$

Attention. Ne pas confondre indépendance et incompatibilité

Exemple 4.1 On lance deux dés et on désigne par A l'événement «le premier dé amène un nombre pair», par B l'événement «le deuxième dé amène un nombre impair» et par C l'événement «les deux dés amènent un nombre pair». On a : $\mathbb{P}(A) = \frac{1}{2}$; $\mathbb{P}(B) = \frac{1}{2}$; $p(C) = \frac{1}{4}$; $\mathbb{P}(A \cap B) = \frac{1}{4}$; $\mathbb{P}(A \cap C) = \frac{1}{2}$; $\mathbb{P}(B \cap C) = 0$. On conclut que A et B sont indépendants ; A et C sont dépendants ; B et C sont dépendants.

Exercice. Soit $\mathcal{F}_n$ l'ensemble des répartitions possibles des sexes dans une famille de n enfants. On suppose toutes les répartitions équiprobables et $n \geq 2$. Les événements M_n : « Avoir des enfants de sexes différents » et F_n : « Avoir au plus une fille » sont-ils indépendants ?

Définition 4.2 n événements $A_1, A_2, \cdots, A_n$ d'un même espace probabilisé sont indépendants (ou indépendants dans leur ensemble) si pour toute partie J de $\{1, 2, \ldots, n\}$ on a

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i) \tag{6}$$

Remarque 4.3 Pour avoir l'indépendance dans l'ensemble, l'indépendance deux à deux ne suffit pas. Voir pour cela un exercice de la fiche d'exercices n°2.

Fin du chapitre 3

Université Hassan II - Mohammadia
F.S.T Mohammadia Département de Mathématiques
Module : Probabilités et statistiques (MIP)

Chapitre 4

Généralités sur les variables aléatoires

1 Introduction

Dans de nombreuses expériences les observations sont des nombres (taille, poids, etc.) D'autre part, lorsque le résultat d'une expérience n'est pas numérique (caractère non quantitatif comme la couleur par exemple), il convient souvent d'y associer un nombre. Cette association est une “variable aléatoire”. Ce chapitre introduit les outils de calcul des probabilités avec les variables aléatoires.

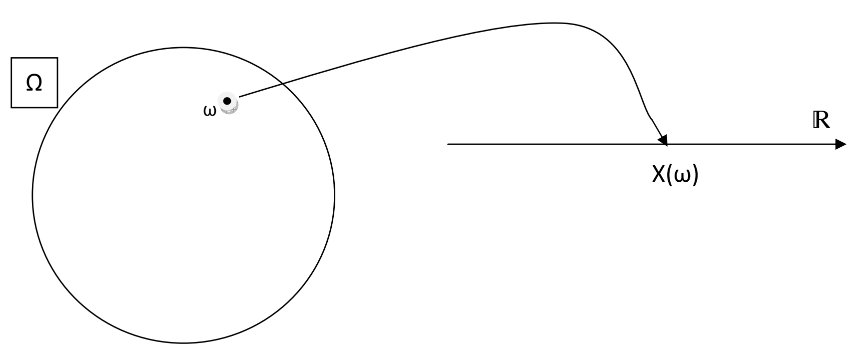
2 Le concept de variable aléatoire

Une variable aléatoire est un nombre réel associé au résultat d'une expérience aléatoire, donc un nombre aléatoire. Si l'épreuve est répétée, ce nombre change en général. Par exemple, la taille d'un individu extrait au hasard d'une population, le nombre de faces dans une série de jets d'une monnaie sont des variables aléatoires. Dans ce qui suit, nous allons formaliser la définition d'une variable aléatoire réelle.

2.1 Définition - Notations

De façon formelle, considérons une épreuve $\mathcal{E}$ et soit Ω l'univers de ses issues. À chaque résultat élémentaire $\omega \in \Omega$, nous attachons un nombre réel $X(\omega) \in \mathbb{R}$. Nous considérons donc une application $X : \Omega \longrightarrow \mathbb{R}$. Une telle application est appelée une *variable aléatoire réelle*. Plus précisément

Définition 2.1 Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une variable aléatoire réelle (en abrégé v.a) est une application $X : \Omega \longrightarrow \mathbb{R}$ telle que pour tout intervalle I de $\mathbb{R}$, l'ensemble $X^{-1}(I) = \{\omega \in \Omega, X(\omega) \in I\}$ soit un événement (il aura donc une probabilité).



Notations conventionnelles En probabilités, l'événement $X^{-1}(I) = \{\omega \in \Omega, X(\omega) \in I\}$ se note $(X \in I)$. Ainsi

$$\begin{array}{lll} X^{-1}(\{x\}) = \{\omega \in \Omega, X(\omega) = x\} & \text{se note} & X = x \\ X^{-1}(]-\infty, b]) = \{\omega \in \Omega, X(\omega) \leq b\} & \text{se note} & X \leq b \\ X^{-1}([a, +\infty[) = \{\omega \in \Omega, X(\omega) \geq a\} & \text{se note} & X \geq a \\ X^{-1}([a, b[) = \{\omega \in \Omega, a \leq X(\omega) < b\} & \text{se note} & a \leq X < b \end{array}$$

Dans la suite nous utiliserons toujours cette notation.

Remarque 2.1 Si la tribu $\mathcal{T}$ est égale à $\mathcal{P}(\Omega)$, ce qui est toujours le cas dans notre cours, toute application de Ω dans $\mathbb{R}$ est une v.a.

Remarque 2.2 L'ensemble image de l'application X est l'ensemble $X(\Omega)$ des valeurs possibles prises par la v.a X.

Examples.

1. On lance une pièce, l'univers associé à cette expérience est $\Omega = \{P, F\}$. Supposons que le fait d'amener Pile soit considéré comme un gain de 1 Dirham et amener Face comme une perte de 1 Dirham, on peut alors considérer $X : \Omega \longrightarrow \mathbb{R}$ définie par $X(P) = 1$ et $X(F) = -1$. Dans ce cas $X(\Omega) = \{-1, 1\}$.
2. Soit une urne à deux catégories contenant des boules blanches (B) et des boules noires (N). On tire dans cette urne n boules avec remise, à chaque tirage ω de n boules on peut faire correspondre le nombre $X(\omega)$ de boules blanches obtenues. Formellement, on peut écrire un résultat élémentaire comme $\omega = (x_1, x_2, \cdots, x_n)$, $x_i = B$ ou $x_i = N$ et la v.a X comme l'application $X \longrightarrow \mathbb{R}$ définie par $X(\omega) =$ nombre d'indices i tels que $x_i = B$. Dans ce cas on a $X(\Omega) = [\![0, n]\!]$.
3. Dans l'urne précédente, on procède à l'expérience suivante : on tire des boules de l'urne avec remise et on s'arrête dès qu'on tire une boule blanche. À chaque tirage ω on fait correspondre le rang $X(\omega)$ où est apparue, pour la première fois, une boule blanche. On dit que la v.a X est le temps d'attente de la première boule blanche. On a alors $X(\Omega) = \mathbb{N}^*$.
4. On joue aux fléchettes. À chaque résultat ω du lancer on fait correspondre la distance $X(\omega)$ du point d'impact au centre de la cible. On obtient ici $X(\Omega) = \mathbb{R}_+$.
5. On lance un oeuf *frais* du haut du twin center et à tout résultat ω on fait correspondre le nombre $X(\omega)$ de rebonds (? ?). On a donc $X(\Omega) = \{0\}$.

3 Fonction de répartition d'une v.a

Définition 3.1 Soit X une v.a d'un espace probabilisé $(\Omega, \mathbb{T}, \mathbb{P})$. On appelle fonction de répartition de X, la fonction numérique F_X définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R}, \quad F_X(x) = \mathbb{P}(X \leq x)$$

On verra plus loin comment la fonction de répartition permet de calculer $\mathbb{P}(X \in I)$ pour tout intervalle I de $\mathbb{R}$.

Exemple 3.1 On lance deux dés et on appelle X la v.a égale à la somme des points obtenus. Donner la représentation graphique de la fonction de répartition de X. On précisera en particulier les points de discontinuité et leur nature.

L'univers associé à cette expérience est $\Omega = [\![1, 6]\!]^2$ et la v.a X est définie par

$$X : \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{N} \\ (i, j) & \longmapsto & i + j \end{array}$$

Donc $X(\Omega) = [\![2, 12]\!]$. Déterminons les probabilités des événements $X = k$ pour $k \in [\![2, 12]\!]$. On a

$$(X = k) = \{(i, j) \in \Omega, i + j = k\}$$

Par exemple l'événement $(X = 7)$ s'écrit $\{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ et par conséquent $\mathbb{P}(X = 7) = \frac{6}{36}$. Pour calculer $F_X(7) = \mathbb{P}(X \leq 7)$, on écrit $(X \leq 7) = (X = 2) \cup (X = 3) \cup (X = 4) \cup (X = 5) \cup (X = 6) \cup (X = 7)$. Donc, $F_X(7) = \sum_{k=2}^{7} \mathbb{P}(X = k)$. En général

Pour $x < 2$, $F_X(x) = 0$

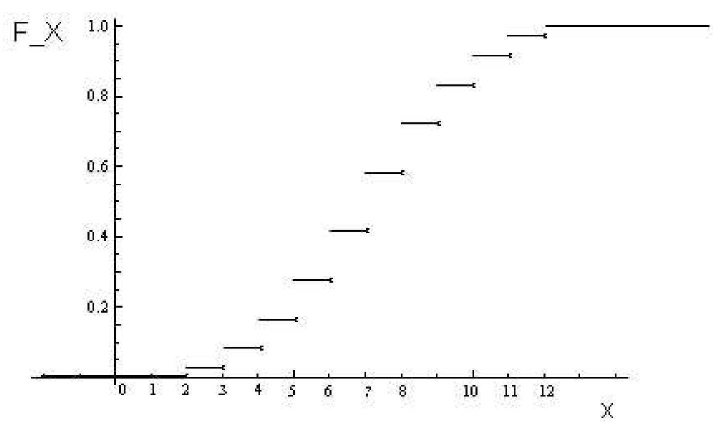
Pour $2 \leq n \leq 11$ et $n \leq x < n + 1$, $F_X(x) = \sum_{k=2}^{n} \mathbb{P}(X = k)$

Pour $x \geq 12$, $F_X(x) = 1$

En effet, $\mathbb{P}(X < 2) = 0$ et pour $n \leq x < n + 1$, $(X \leq x) = \bigcup_{i=2}^{n}(X = i)$, d'où $F_X(x) = \sum_{k=2}^{n} \mathbb{P}(X = k)$ et finalement, l'événement $X \geq 12$ est certaine et par conséquent $F_X(x) = 1$:

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
$F_X(k) = \sum_{i=2}^{k} \mathbb{P}(X = i)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	1

On vérifie bien que $\sum_{k=1}^{12} \mathbb{P}(X = k) = 1$. Ci-dessous, est représenté le graphe de F_X.



La fonction de répartition est donc une fonction en escalier, croissante, continue à droite.

En général, on remarque que si $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$ avec $x_1 < x_2 < \cdots < x_n$, on a

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < x_1 \\ 1 & \text{si} \quad x \geq x_n \\ F_X(x_k) = \sum_{i=1}^{k} \mathbb{P}(X = x_i) & \text{si} \quad x_k \leq x < x_{k+1} \end{cases}$$

3.1 Propriétés d'une fonction de répartition

La fonction de répartition d'une variable aléatoire admet 3 propriétés fondamentales résumées dans le théorème suivant que nous admettrons

Théorème 3.2 Soit X une v.a et soit F_X sa fonction de répartition. Alors F_X possède les propriétés suivantes

1. F_X est une fonction croissante
2. F_X est continue à droite en tout point de $\mathbb{R}$
3. $\lim_{x \to -\infty} F_X(x) = 0$ et $\lim_{x \to +\infty} F_X(x) = 1$

Nous admettrons aussi que toute fonction numérique $F : \mathbb{R} \longrightarrow [0,1]$ vérifiant les trois propriétés précédentes est la fonction de répartition d'une certaine v.a X.

Exemple 3.2 On considère la fonction F définie sur $\mathbb{R}$ par

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{x}{8} & \text{si} \quad 0 \leq x < 2 \\ \frac{x}{5} & \text{si} \quad 2 \leq x < 4 \\ 1 & \text{si} \quad x \geq 4 \end{cases}$$

Tracer la représentation graphique de F et montrer que F peut être considérée comme la fonction de répartition d'une v.a X.

D'après le graphe de F, on voit bien que F vérifie les trois propriétés du théorème. À savoir,

- F est croissante au sens large,
- F est continue sur $\mathbb{R}$ sauf aux points $x = 2$ et $x = 4$ en lesquels, elle est continue à droite. En effet $F(2) = 2/5$ et $F(4) = 1$ par définition et

$$\lim_{x \to 2^-} = \frac{1}{4} \quad \text{et} \quad \lim_{x \to 2^+} = \frac{2}{5} = F(2)$$

De même

$$\lim_{x \to 4^-} = \frac{4}{5} \quad \text{et} \quad \lim_{x \to 4^+} = 1 = F(4)$$

- Évidement, la limite en $-\infty$ de F est nulle et celle en $+\infty$ est égale à 1.

On en déduit que la fonction numérique F peut être considérée comme la fonction de répartition d'une v.a réelle.

La proposition suivante montre comment calculer $\mathbb{P}(X \in I)$ pour un intervalle I de $\mathbb{R}$ à partir de la fonction de répartition

Proposition 3.1 Soit X une v.a et F_X sa fonction de répartition. Alors

1. $\forall (a,b) \in \mathbb{R}^2$ avec $a < b$, $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$
2. $\forall a \in \mathbb{R}$, $\mathbb{P}(X = a) = F_X(a) - \lim_{x \to a^-} F_X(x)$

Remarque 3.1 Si F_X est continue en a alors $\lim_{x \to a^-} F(x) = F(a)$, ce qui entraîne $\mathbb{P}(X = a) = 0$

Les deux formules de la proposition permettent de calculer $\mathbb{P}(X \in I)$ pour tout intervalle de $\mathbb{R}$.
• $I = [a, b]$: $\mathbb{P}(a \leq X \leq b) = \mathbb{P}(X = a) + \mathbb{P}(a < X \leq b) = \mathbb{P}(X = a) + F_X(b) - F_X(a)$
• $I =]a, b[$: $\mathbb{P}(a < X < b) = \mathbb{P}(a < X \leq b) - \mathbb{P}(X = b) = F_X(b) - F_X(a) - \mathbb{P}(X = b)$
• $I = [a, +\infty[$: $\mathbb{P}(X \geq a) = 1 - \mathbb{P}(X < a) = 1 - \mathbb{P}(X \leq a) + \mathbb{P}(X = a) = 1 - F_X(a) + \mathbb{P}(X = a)$
etc.

Exercice . Soit F la fonction définie par :

$$\begin{cases} F(x) = 0, & \text{si} \quad x < 0 \\ F(x) = \frac{x}{2}, & \text{si} \quad 0 \leq x < 1 \\ F(x) = \frac{2}{3}, & \text{si} \quad 1 \leq x < 2 \\ F(x) = \frac{11}{12}, & \text{si} \quad 2 \leq x < 3 \\ F(x) = 1, & \text{si} \quad x \geq 3 \end{cases}$$

Montrer que F est une fonction de répartition d'une v.a. X.
Calculer $\mathbb{P}(X = 1/2)$; $\mathbb{P}(X = 1)$; $\mathbb{P}(X < 3)$; $\mathbb{P}(2 < X \leq 4)$.

Fin du chapitre 4

Chapitre 5

Variables aléatoires discrètes

Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. On dit qu'une variable aléatoire réelle est discrète (v.a.d en abrégé) si l'ensemble $X(\Omega)$ de ses valeurs prises est dénombrable.

1 Loi d'une v.a.d

Définition 1.1 Soit X une v.a.d. L'application f de $X(\Omega)$ dans $\mathbb{R}$ qui à tout $x \in X(\Omega)$ fait correspondre $f(x) = \mathbb{P}(X = x)$ s'appelle la *loi de probabilité* de X (on dit aussi *distribution*).

Pour déterminer la loi de X, on commence toujours par déterminer Ω et $X(\Omega)$.

Remarque 1.1 Si X est une v.a.d et f sa loi, alors

$$\sum_{x \in X(\Omega)} f(x) = 1$$

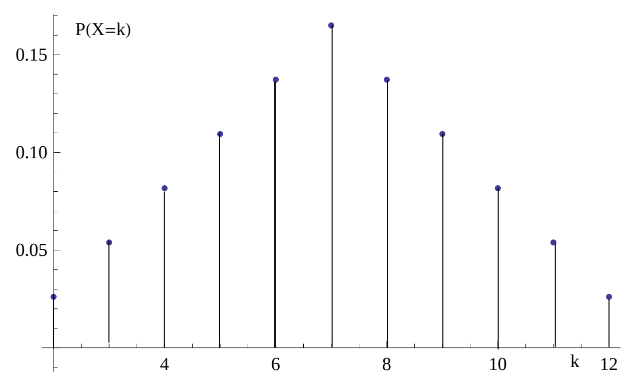
Car on peut écrire Ω comme la réunion disjointe et dénombrable des événements $(X = x)$, $x \in X(\Omega)$. Le résultat s'ensuit puisque $\mathbb{P}$ est une probabilité.

On démontre que la réciproque de ce résultat est aussi valable.

Proposition 1.1 *Soit X une v.a.d. La fonction de répartition F_X est entièrement déterminée par la loi de X.*

Il suffit donc dans le cas discret de calculer la loi de X. L'exemple 3.1 du chapitre 4 est un exemple de v.a.d.

Pour visualiser graphiquement la loi d'une v.a.d, on trace un diagramme en bâtons qui consiste à représenter dans un repère les segments d'extrémités les points de coordonnées $(x, 0)$ et $(x, \mathbb{P}(X = x))$ pour $x \in X(\Omega)$. Par exemple la loi de l'exemple 2.1 du chapitre 4 peut être visualisée par le diagramme en bâtons suivant.



2 Espérance, variance, écart-type

Pour une v.a.d X, nous posons $X(\Omega) = \{x_1, x_2, \ldots, x_n, \ldots\}$ ($X(\Omega)$ est dénombrable par définition). Pour définir l'espérance mathématique, ou moyenne des valeurs prises par la v.a.d, nous distinguons deux cas : $X(\Omega)$ fini et infini.

Définition 2.1 Si $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$ est fini, on définit l'espérance de X, que l'on note $\mathbb{E}(X)$, par la somme finie suivante :

$$\mathbb{E}(X) = \sum_{i=1}^{n} x_i \mathbb{P}(X = x_i)$$

Si $X(\Omega)$ est infinie, l'espérance existe si et seulement si la série $\sum_i x_i \mathbb{P}(X = x_i)$ est absolument convergente, c'est à dire que la somme $\sum_i |x_i| \mathbb{P}(X = x_i)$ est finie et l'on a dans ce cas

$$\mathbb{E}(X) = \sum_{i=1}^{\infty} x_i \mathbb{P}(X = x_i)$$

Exemple 2.1 On joue pile ou face. L'apparition de Pile fait gagner un Dirham et celle de Face fait perdre un Dirham. Quelle est l'espérance du gain ?

Ici le gain est une v.a.d X prenant les valeurs $+1$ et -1 avec la probabilité $1/2$ pour chaque face. C'est à dire c'est l'application de $\Omega = \{P, F\}$ dans $\{-1, +1\}$ définie par $X(\{P\}) = +1$ et $X(\{F\}) = -1$. On a $\mathbb{P}(X = -1) = \mathbb{P}(X = +1) = 1/2$. L'espérance de X est donc $\mathbb{E}(X) = (+1)\frac{1}{2} + (-1)\frac{1}{2} = 0$. En moyenne, on n'espère rien gagner.

2.1 Moments d'une v.a.d

Définition 2.2 On appelle moment d'ordre k de la v.a.d X la quantité

$$\mu_k(X) = \sum_{x \in X(\Omega)} x^k \mathbb{P}(X = x)$$

Ainsi l'espérance n'est autre que le moment d'ordre 1.

Le moment d'ordre k de la v.a.d de l'exemple 2.1 est $\left((-1)^k + 1\right)/2$ égal à 1 si k est pair et à 0 si k est impair.

2.2 Variance et écart-type

Pour introduire la notion de variance, discutons un exemple simple. Considérons deux v.a.d X et Y dont les lois sont données par

x	95	97	100	103	105
$\mathbb{P}(X = x)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

et

y	50	75	100	125	150
$\mathbb{P}(Y = y)$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

Les deux v.a, dont les diagrammes en bâtons sont représentés ci-dessous, ont la même espérance (la même moyenne) $\mathbb{E}(X) = \mathbb{E}(Y) = 100$, mais on voit que la loi de Y est plus “dispersée” que celle de X.

La question qui se pose est de définir un paramètre qui caractérise cette dispersion. Ce paramètre est l'écart-type par lequel on mesure l'écart de la loi par rapport à la moyenne (qui est le "centre" de la loi).

Définition 2.3 Si l'espérance d'une v.a.d X existe, la variance de X est, si elle existe, la somme

$$\mathrm{Var}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x)$$

L'écart-type de X, noté σ_X, est la racine carrée de la variance.

$$\sigma(X) = \sqrt{\mathrm{Var}(X)}$$

2.3 Propriétés de l'espérance

Proposition 2.1

1. *Linéarité : Soient X et Y deux v.a.d définies sur un même espace probabilisé $(\Omega, \mathcal{T}, \mathbb{P})$ et ayant chacune une espérance. Soit $\alpha \in \mathbf{R}$ alors $X + Y$ et αX ont une espérance et on a :*

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y) \quad et \quad \mathbb{E}(\alpha X) = \alpha \mathbb{E}(X)$$

2. *Soit ϕ une application de $\mathbf{R}$ dans $\mathbf{R}$, la v.a.d $\phi(X)$ a une espérance si et seulement si la somme $\sum_{x \in X(\Omega)} \phi(x)\mathbb{P}(X = x)$ existe et on a alors*

$$\mathbb{E}(\phi(X)) = \sum_{x \in X(\Omega)} \phi(x)\mathbb{P}(X = x)$$

Remarque 2.1 Le calcul de l'espérance de la v.a.d $\phi(X)$ doit se faire à partir de sa loi

$$\mathbb{E}(\phi(X)) = \sum_{x \in \phi(X)(\Omega)} y\mathbb{P}(\phi(X) = y)$$

La propriété 2 montre qu'il est inutile de déterminer la loi de $\phi(X)$ pour calculer son espérance.

Comme application de la propriété 2, on peut écrire

$$\mathrm{Var}(X) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \, \mathbb{P}(X = x) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$$

donc

$$\mathrm{Var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$$

Cette formule est souvent prise comme définition de la variance. Ainsi la variance mesure la moyenne des carrés des écarts à la moyenne.

2.4 Propriétés de la variance

Il est facile de démontrer la

Proposition 2.2 (Koenig-Huyghens) *Soit X une v.a.d, ayant une espérance et une variance, alors la v.a.d X^2 a une espérance et*

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Ce théorème donne une méthode pratique de calcul de la variance.

Proposition 2.3 *Soit X une v.a.d ayant une espérance et une variance alors pour tout $(a, b) \in \mathbf{R}^2$, la variable $aX + b$ a une variance et*

$$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$$

Preuve : Exercice facile.

Définition 2.4 Soit X une v.a.d ayant une espérance et une variance. La v.a

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$$

est appelée variable aléatoire centrée réduite associée à la v.a X. D'après la proposition précédente, elle est telle que son espérance est nulle et sa variance est égale à 1.

3 Quelques lois discrètes classiques

3.1 Variable certaine

Par définition, c'est une variable constante :

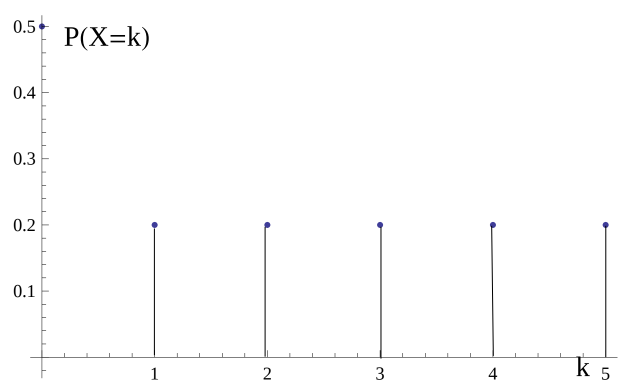
$$X(\Omega) = \{a\}, \quad \mathbb{P}(X = a) = 1, \quad \mathbb{E}(X) = a, \quad \mathrm{Var}(X) = 0$$

3.2 Loi uniforme discrète sur ⟦1,n⟧

Définition 3.1 On dit qu'une v.a X suit une *loi uniforme discrète* sur $[\![1, n]\!]$ si l'on a $X(\Omega) = [\![1, n]\!]$ et

$$\forall k \in [\![1, n]\!] \quad \mathbb{P}(X = k) = \frac{1}{n}$$

Modèle. On tire au hasard une boule dans une urne contenant n boules numérotées de 1 à n ; X désigne le numéro de la boule tirée.



Remarque 3.1 Pour $n = 1$, X suit une loi certaine.

Proposition 3.1 *Soit X une v.a.d uniforme sur $[\![1, n]\!]$ alors*

$$\mathbb{E}(X) = \frac{n+1}{2} \qquad \mathrm{Var}(X) = \frac{n^2-1}{12}$$

Preuve Si X est une v.a.d uniforme sur $[\![1, n]\!]$. On a

$$\mathbb{E}(X) = \sum_{k=1}^{n} \frac{k}{n} = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$$

et

$$\mathbb{E}(X^2) = \sum_{k=1}^{n} \frac{k^2}{n} = \frac{1}{n}\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

où l'on a utilisé la formule qui donne la somme des carrés $\sum_{1\leq k\leq n} k^2$ (que l'on peut démontrer par récurrence). Il vient alors

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{n^2-1}{12}$$

□

3.3 Loi de Bernoulli

Définition 3.2 Une v.a.d suit une loi de Bernoulli de paramètre p si elle prend les valeurs 1 et 0 avec les probabilités p et $q = 1 - p$ respectivement, où $0 < p < 1$. p est le paramètre caractérisant la loi de Bernoulli. On dit aussi que X est une v.a de Bernoulli de paramètre p.

Modèle : On considère une expérience aléatoire ayant deux résultats que nous appelons succès (noté S) et échec (noté E) avec les probabilités p et $q = 1-p$. L'univers des possibilités est donc $\Omega = \{S, E\}$. La variable de Bernoulli est définie par $X(S) = 1$ et $X(E) = 0$ avec les probabilités $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = q$

Proposition 3.2 *Soit X une v.a de Bernoulli alors*

$$\mathbb{E}(X) = p \quad \text{et} \quad \mathrm{Var}(X) = pq$$

Preuve : On a $\mathbb{E}(X) = p \times 1 + q \times 0 = p$ et on a aussi $\mathbb{E}(X^2) = p$. Donc $\mathrm{Var}(X) = p - p^2 = pq$. □

3.4 Loi binomiale

Définition 3.3 Une v.a.d suit une loi binomiale de paramètres n et p ($n \in \mathbf{N}^*$ et $0 < p < 1$) si $X(\Omega) = [\![0, n]\!]$ et si pour tout $k \in [\![0, n]\!]$, on a

$$\mathbb{P}(X = k) = C_n^k p^k q^{n-k}, \quad (q = 1 - p)$$

On dit aussi que X est une variable Binomiale de paramètres n et p.

On peut s'assurer de la cohérence de cette définition en remarquant que

$$\sum_{k=0}^{n} \mathbb{P}(X = k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1$$

Comment reconnaître une loi binomiale ? On considère une expérience aléatoire à deux issues {Succès, Echec}. La probabilité du Succès étant égale à p. On répète cette expérience n fois dans les mêmes conditions et on s'intéresse au nombre de réalisations du succès. Soit X la v.a.d qui compte ce nombre. Alors X est une v.a binomiale de paramètres n et p.

Exemple 3.1 Reconnaître une v.a de loi binomiale dans chacune des situations suivantes.

1. On lance 4 fois deux dés. Quelle est la loi de la v.a égale au nombre d'apparitions de la somme 4 ? Quelle est la probabilité d'obtenir au moins 2 fois une somme égale à 4 ?
2. On dispose d'un lot de 100 articles dont cinq sont défectueux. On tire au hasard et avec remise 20 articles. Quelle est la probabilité qu'ils soient tous sans défaut ?

Proposition 3.3 *Soit X une v.a de loi binomiale de paramètres n et p. Alors*

$$\mathbb{E}(X) = np \quad \mathrm{var}(X) = npq$$

Exercice Montrer la proposition 3.3.

Proposition 3.4 *Posons $p_n = \frac{\lambda}{n}$. Alors pour k entier fixé. On a*

$$\lim_{\substack{n \to +\infty \\ n > k}} C_n^k p_n^k (1 - p_n)^{n-k} = e^{-\lambda} \frac{\lambda^k}{k!}$$

Preuve. exercice.

Cette proposition donne une possibilité pratique pour calculer numériquement la probabilité de la loi binomiale si le paramètre n est grand. En général, on admet dans la pratique que pour $n \geq 50$, $np \leq 15$, $p \leq 0.1$, la valeur de $C_n^k p^k q^{n-k}$ est bien approchée par la valeur $e^{-\lambda} \frac{\lambda^k}{k!}$ avec $\lambda = np$.

3.5 Loi de Poisson

Définition 3.4 On dit qu'une v.a.d X suit une loi de Poisson de paramètre $\lambda > 0$ si $X(\Omega) = \mathbf{N}$ et

$$\forall k \in \mathbf{N} \qquad \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

On a bien une probabilité car

$$\sum_{n=0}^{\infty} \mathbb{P}(X = n) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

En pratique, la loi de Poisson modélise les expériences dont les événements sont rares. Par exemple, en physique nucléaire, le nombre de particules α émises par une substance radioactive dans un intervalle de temps Δt est une v.a de Poisson de paramètre proportionnel à Δt.

Proposition 3.5 *Une v.a suivant une loi de Poisson de paramètre λ admet une espérance et une variance données par*

$$\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$$

Preuve : Voir le cours.

Exemple 3.2 .

1. On admet que le nombre d'appels téléphoniques reçus par un standard, durant une période de temps T (exprimée en heures) suit une loi de Poisson de paramètre $\lambda = 10T$. Donner la probabilité que le nombre d'appels reçus dans une période de 6 min soit égale ou supérieur à 4.
2. On admet que le nombre d'étoiles filantes observées durant une période de T minutes, à une époque donnée, suit une loi de Poisson de paramètre $\lambda = aT$ (a étant une constante indépendante de T). Déterminer a sachant qu'on observe en moyenne une étoile filantes toutes les 2 minutes. Quelle est la probabilité d'observer au moins 5 étoiles filantes en 10 minutes ?

3.6 Loi géométrique

Modèle : On répète n fois une expérience aléatoire à deux issues (à la Bernoulli) jusqu'à l'obtention du premier succès. X est la v.a qui compte le rang du premier succès :

$$\Omega = \{\omega_n = (\underbrace{EEE\cdots E}_{n-1 \text{ fois}} S), n \in \mathbf{N}^*\}$$

La v.a X est donc l'application qui fait correspondre au résultat élémentaire ω_n l'indice n qui constitue le "temps d'attente" de l'apparition du premier succès, comme un auto-stoppeur qui attend la voiture qui le transporte. Il est immédiat de voir que $\mathbb{P}(X = n) = \omega_\ltimes = pq^{n-1}$, $(q = 1 - p)$. D'où la

Définition 3.5 On dit qu'une v.a X suit une loi géométrique de paramètre p $(0 < p < 1)$ si $X(\Omega) = \mathbf{N}^*$ et

$$\forall k \in \mathbf{N}^*, \quad \mathbb{P}(X = k) = pq^{n-1}, \quad q = 1 - p$$

La définition est cohérente car

$$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = p \sum_{k=1}^{\infty} q^{n-1} = p \sum_{i=0}^{\infty} q^i = \frac{p}{1-q} = 1$$

où l'on a appliqué la formule donnant la somme d'une série géométrique. Pour cette raison, cette loi s'appelle géométrique. On l'appelle aussi loi de l'auto-stoppeur.

Proposition 3.6 *L'espérance et la variance d'une v.a de loi géométrique de paramètre p sont donnés par*

$$\mathbb{E}(X) = \frac{1}{p} \qquad \mathrm{Var}(X) = \frac{q}{p^2}$$

Preuve. Voir le cours.

3.7 Loi hypergéométrique

Modèle : Une urne contient N_1 boules blanches et N_2 boules noires. On fait un tirage de n boules simultanément. Quelle est la loi de la variable aléatoire X qui compte le nombre de boules blanches dans l'échantillon ?

Ici le nombre total des possibilités est $\mathrm{Card}(\Omega) = C^n_{N_1+N_2}$. Soit l'événement « Il y a k boules blanches dans l'échantillon » $= (X = k)$. On a $\mathrm{Card}(X = k) = C^k_{N_1} C^{n-k}_{N_2}$, c'est le nombre de façons

de choisir k boules parmi N_1 boules blanches multiplié par le nombre de choix de $n-k$ boules parmi N_2 boules noires (principe de multiplication). D'où

$$\mathbb{P}(X = k) = \frac{C_{N_1}^k C_{N_2}^{n-k}}{C_{N_1+N_2}^n}$$

On dit que la v.a X suit une loi hypergéométrique de paramètres N_1, N_2 et n.

L'ensemble des valeurs de X dépend des valeurs de ses paramètres. On monte qu'en général

$$X(\Omega) = \{\max(N_1, n - N_2), \ldots, \min(N_1, n - N_2)\}$$

Proposition 3.7 *L'espérance et la variance d'une v.a suivant une loi hypergéométrique sont données par*

$$\mathbb{E}(X) = np \qquad \mathrm{Var}(X) = npq\frac{N_1 + N_2 - n}{N_1 + N_2 - 1}$$

où

$$p = \frac{N_1}{N_1 + N_2} \quad q = 1 - p$$

Fin du chapitre 5

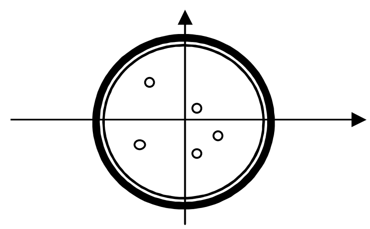
Chapitre 6

Variables aléatoires continues

Dans le chapitre précédent, on s'est limité aux variables aléatoires discrètes (qui ne peuvent prendre qu'un nombre fini ou infini dénombrable de valeurs). Mais il existe beaucoup de variables aléatoires qui peuvent prendre n'importe quelle valeur de $\mathbb{R}$ ou d'un intervalle de $\mathbb{R}$.

1. Exemple préliminaire

L'exemple que nous allons considérer introduit d'une manière intuitive la notion de densité de probabilité.



On lance une fléchette sur une cible matérialisée par un disque de rayon 1, et on mesure la distance R du point d'impact par rapport au centre du disque.

En supposant que la fléchette atteint toujours la cible, on peut dire que notre jeu consiste à choisir au hasard un point dans le disque centré à l'origine du repère et de rayon 1. Il est donc clair que la variable aléatoire R est à valeurs dans l'intervalle $[0,1]$ qui n'est pas dénombrable.

Commençons par poser la question suivante : « Quelle est la probabilité que le point d'impact soit sur un cercle de rayon π ?, de rayon $\pi/4$?". Sans entrer dans les détails de ce qu'on entend par « au hasard », il paraît évident que la probabilité que le point d'impact se situe à une distance égale à π par rapport à l'origine est nulle $(\pi > 1)$. La probabilité que la fléchette « marque » un point à $\pi/4$ de l'origine est également nulle, non pas parce que l'événement est impossible, mais du fait qu'il correspond à un nombre irrationnel $(\pi/4 \cong 0.7853981635)$ et qu'il serait nécessaire d'obtenir la coïncidence d'une infinité de décimales.[1]

Les événements $(R = \pi)$ et $(R = \pi/4)$ sont donc de probabilité nulle, mais pour deux raisons différentes : la probabilité que le point d'impact soit « approximativement » distant de π par rapport à l'origine est nulle et la probabilité qu'il soit « approximativement » distant de $\pi/4$ par rapport à l'origine n'est pas nulle. On dira qu'il existe au voisinage des points du cercle de rayon $\pi/4$ une densité de probabilité non nulle.

[1] Intuitivement, le point d'impact sera peut-être proche d'un point se trouvant sur un cercle de rayon $\pi/4$, mais en regardant avec une loupe, la pointe de la fléchette sera presque sûrement dans un voisinage de ce point.

2. Densité et fonction de répartition

Définition 2.1. Soit X une v.a. et F_X sa fonction de répartition. On dira (dans le cadre de ce cours) que X est une v.a. continue s'il existe une fonction $f \geq 0$ telle que pour tout $x \in \mathbb{R}$, l'intégrale $\int_{-\infty}^{x} f(t)\,dt$ ait un sens et si l'on a

$$\forall x \in \mathbb{R},\ \ F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)\,dt \qquad (1)$$

La fonction f s'appelle une **densité de probabilité** de X.

Tous les calculs de probabilités relatifs à X peuvent être traités à l'aide de la formule (1), comme le montre la proposition suivante qui découle des propriétés des intégrales et des variables aléatoires.

Proposition 2.1. *Soit X une v.a continue, F_X sa fonction de répartition et f sa densité.*

1. *La fonction de répartition F_X est continue sur $\mathbb{R}$.*
2. $\forall a \in \mathbb{R},\ \mathbb{P}(X = a) = 0$ *et* $\forall (a, b) \in \mathbb{R}^2$ *tels que*

$$\begin{aligned} F_X(b) - F_X(a) &= \int_a^b f(t)\,dt = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X < b) \\ &= \mathbb{P}(a \leq X < b). \end{aligned} \qquad (2)$$

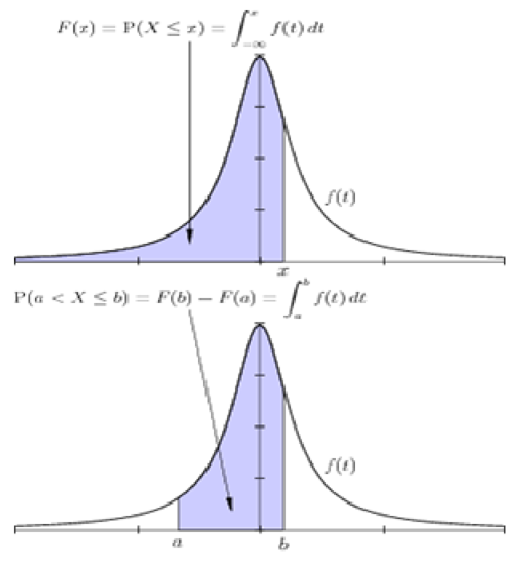
3. *La fonction de répartition est continûment dérivable sur tout intervalle de $\mathbb{R}$ sur lequel la dérivée f est continue. En tout point x où f est continue, on a*

$$f(x) = F_X'(x).$$

2.1. Interprétation de la densité. En tout point $a \in \mathbb{R}$ où f est continue, la formule $f(a) = F_X'(a)$ peut s'écrire

$$f(a) = \lim_{b \to a^+} \frac{F_X(b) - F_X(a)}{b - a} = \lim_{b \to a^+} \frac{\mathbb{P}(a < X \leq b)}{b - a}$$

$f(a)$ correspond bien à la densité de probabilité en a et la valeur $\mathbb{P}(a < X \leq b)$ peut être réalisé par l'aire limitée par la courbe d'équation $y = f(x)$ et les droites d'équations $y = 0$, $x = a$ et $x = b$.



2.2. Propriétés de la densité

Proposition 2.2. *Une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ est densité de probabilité d'une v.a. continue si*

1. $f(x) \geq 0 \quad \forall x \in \mathbb{R}$,
2. *L'intégrale $\int_{-\infty}^{+\infty} f(t)\,dt$ converge et est égale à 1.*

Exemple 2.1. Supposons qu'une v.a. X a la densité f telle que:

$$f(x) = \begin{cases} C(4x - 2x^2), & 0 < x < 2 \\ 0\,, & \text{Sinon} \end{cases}$$

a) Quelle est la valeur de C? Que vaut $\mathbb{P}(X > 0)$ et $\mathbb{P}(X > 1)$?
b) Déterminer et représenter la fonction de répartition de X.

Exemple 2.2. La durée de fonctionnement d'un ordinateur avant sa première panne est une v.a. continue de densité donnée par

$$f(x) = \begin{cases} \lambda\, e^{-x/100}, & x > 0 \\ 0\,, & \text{Sinon} \end{cases}$$

a) Quelle est la probabilité que cette durée soit comprise entre 50 et 150 jours?
b) Quelle est la probabilité que l'ordinateur fonctionne moins de 100 jours?
c) Déterminer et représenter la fonction de répartition de cette v.a.

2.3. Exemples de détermination de la densité d'une fonction de v.a continue. On se pose la question suivante : « X étant une v.a. continue et ϕ une application de $\mathbb{R}$ dans $\mathbb{R}$. La v.a. $\phi(X)$ est-elle encore une v.a. continue? Si oui, comment déterminer sa densité en fonction de celle de X ? ». La proposition suivante répond à cette question quand $\phi(t) =$constante, $\phi(t) = at + b$ et $\phi(t) = t^2$.

Proposition 2.3. *Soit F_X la fonction de répartition de la v.a. X continûment dérivable sauf peut-être en un nombre fini de points $\{a_1, a_2, \dots, a_n\}$.*

i. *Si ϕ est une application constante $\phi(X) = c$ alors*

$$F_{\phi(X)}(x) = \begin{cases} 0, & x < c \\ 1, & x \geq c \end{cases}$$

Dans ce cas $\phi(X)$ n'est pas une v.a. continue.

ii. *Si ϕ est une application affine $\phi(X) = aX + b$ $(a \neq 0)$ alors $\phi(X)$* est une v.a. *continue de densité*

$$\forall x \in \mathbb{R}, \qquad f_{\phi(X)}(t) = \frac{1}{|a|} f\left(\frac{t-b}{a}\right)$$

iii. *Si $\phi(X) = X^2$ alors $\phi(X)$ est une v.a. continue de densité*

$$f_{X^2}(t) = \begin{cases} 0, & t < 0 \\ \dfrac{1}{2\sqrt{t}}\Big(f_X(\sqrt{t}) + f_X(\sqrt{t})\Big), & t > 0 \ \text{ et } \ t \notin \{a_1^2, \dots, a_n^2\} \end{cases}$$

Preuve. Voir cours. ■

3. Espérance, variance d'une v.a. continue

Par analogie avec les variables aléatoires discrètes, on définit l'espérance d'une v.a. continu de la façon suivante

Définition 3.1. Soit X une v.a. continue de densité f. On appelle espérance de X l'intégrale (si elle existe)

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} t\, f(t)\, dt \quad (3)$$

Si l'espérance de X existe, la variance de X est définie par l'intégrale (si elle existe)

$$\mathrm{Var}(X) = \int_{-\infty}^{+\infty} (t - \mathbb{E}(X))^2\, f(t)\, dt \quad (4)$$

L'écart-type de X est

$$\sigma(X) = \sqrt{\mathrm{Var}(X)} \quad (5)$$

3.1. Espérance de $Y = \phi(X)$. Nous admettons le résultat important suivant

Proposition 3.1. *Soit X une v.a. de densité f et soit $\phi: \mathbb{R} \to \mathbb{R}$ une application **continue ou continue par morceaux**. Si l'intégrale $\int_{-\infty}^{+\infty} |\phi(t)| f(t)\, dt$ existe, alors la v.a $\phi(X)$ a une espérance donnée par*

$$\mathbb{E}(\phi(X)) = \int_{-\infty}^{+\infty} \phi(t) f(t)\, dt \quad (6)$$

3.2. Quelques propriétés. Il est très facile de prouver les propriétés remarquables suivantes

Proposition 3.2.

- *Si X est une v.a. continue ayant une espérance alors la v.a. $aX + b$ admet une espérance et $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.*
- *Si X a une variance alors la v.a. $aX + b$ a une variance et $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$.*
- ***Théorème de Koenig-Huyghens**. Si X a une espérance alors elle a une variance si et seulement si la v.a. X^2 a une espérance et dans ce cas, on a*

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

Exemple 3.1. Calculer les espérances et les variances des v.a. des exemples 2.1 et 2.2.

4. Lois continues usuelles.

Les v.a continues classiques sont toujours définies à partir d'une densité.

4.1. Loi uniforme

La loi uniforme est utilisée pour modéliser une grandeur aléatoire qui prend au hasard (c'est à dire, sans préférence pour aucune d'elles) ses valeurs dans un intervalle.

Définition 4.1. Une v.a. X est une v.a. uniforme sur l'intervalle $[a, b]$ si sa densité est donnée par

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0\,, & \text{Sinon} \end{cases} \qquad (7)$$

On note $X \rightsquigarrow \mathcal{U}([a,b])$.

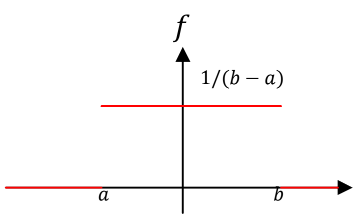
Il est évident que f vérifie les propriétés 1 et 2 de la proposition 2.2.

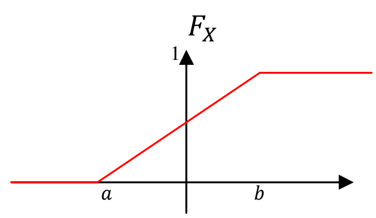
La fonction de répartition d'une v.a suivant une loi uniforme sur l'intervalle $[a,b]$ est $F_X(x) = \int_{-\infty}^{x} f(t)\,dt$. Donc

si $x < a$ $\qquad F_X(x) = 0$

si $a \leq x \leq b$ $\qquad F_X(x) = \int_{-\infty}^{a} 0\,dt + \int_{a}^{x} f(t)\,dt = \dfrac{x-a}{b-a}$

si $x \geq b$ $\qquad F_X(x) = 1$





Exemple 4.1. A partir de 7 heures, les bus passent toutes les 15 minutes à un arrêt donné. Un usager se présente entre 7 et 7, 30, l'heure de son arrivée étant une v.a. uniforme sur cette période. Trouver la probabilité pour qu'il doit attendre moins de 5 minutes; plus de 10 minutes.

Réponse : $\frac{1}{6}, \frac{2}{3}$.

Exemple 4.2. Un point est choisi au hasard sur un segment de longueur L. Interpréter cet énoncé et trouver la probabilité pour que le rapport entre le plus petit et le plus grand segment soit inférieure à 1/4.

Réponse : $\frac{2}{5}$.

Proposition 4.1. *Soit $X \rightsquigarrow \mathcal{U}([a,b])$, alors X a une espérance et une variance données par*

$$\mathbb{E}(X) = \frac{a+b}{2}; \quad \mathrm{Var}(X) = \frac{(b-a)^2}{12} \qquad (8)$$

Preuve. Calcul simple. ■

Exemple 4.3. Un nombre choisi au hasard dans l'intervalle $[0,1]$ est une v.a uniforme sur $[0,1]$. En moyenne, le nombre choisit se trouve au centre de l'intervalle ($\mathbb{E}(X) = 1/2$) avec une dispersion de $1/\sqrt{12} \cong 0.288675$.

Remarque. Vous pouvez « vérifier » le résultat de l'exemple précédent en appuyant un grand nombre de fois (disons 100 fois) sur la touche de votre calculatrice qui génère un nombre aléatoire entre 0 et 1 (sur certaine calculette, CASIO par exemple, cette touche s'affiche comme 'Ran#') et noter à chaque fois le nombre généré, pour enfin calculer la moyenne de tous les nombres. Vous allez trouver que cette moyenne est très proche de 0.5 !

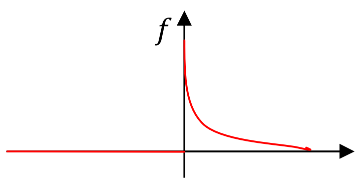
4.2. Loi exponentielle

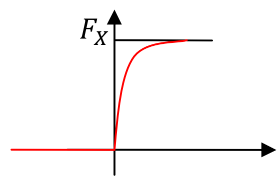
La loi exponentielle est utilisée pour mesurer des temps d'attente, des temps de service etc. On la rencontre souvent en fiabilité pour modéliser des durées de vie des composants.

Définition 4.2. Une v.a. X suit une loi exponentielle de paramètre λ si sa densité est donnée par

$$f(x) = \begin{cases} \lambda\, e^{-\lambda x}, & x \geq 0 \\ 0\,, & \text{Sinon} \end{cases} \tag{9}$$

On note $X \rightsquigarrow \mathcal{E}(\lambda)$.





On trouve facilement

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0\,, & \text{Sinon} \end{cases}$$

et

$$\mathbb{E}(X) = \frac{1}{\lambda}; \quad \mathrm{Var}(X) = \frac{1}{\lambda^2} \tag{10}$$

Exercice. Soit $X \rightsquigarrow \mathcal{E}(\lambda)$. Montrer l'égalité suivante

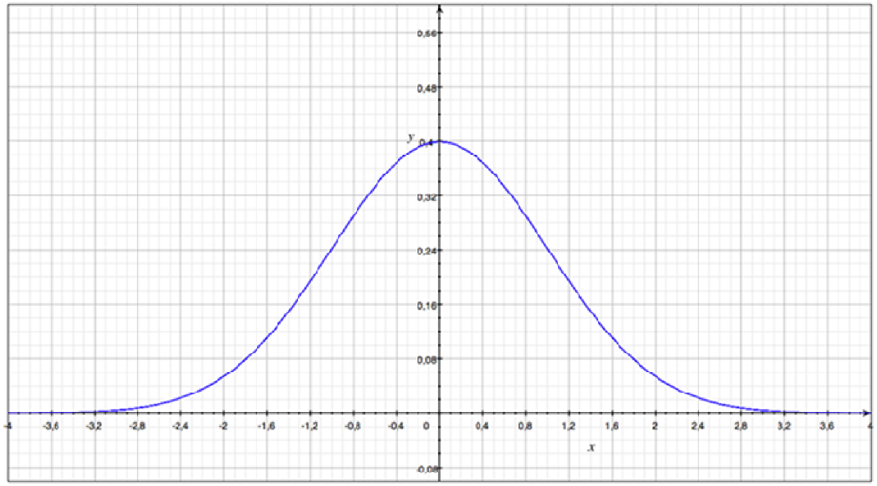
$$\mathbb{P}(X > t + s\,|X > t) = \mathbb{P}(X > s)$$

Cette relation montre qu'une variable aléatoire exponentielle n'a pas de mémoire. Cette absence de mémoire est la propriété la plus importante de la loi exponentielle.

Exemple 4.3. Le temps nécessaire pour réparer une machine est une v.a. exponentielle de moyenne 2 (heures).

a) Quelle est la probabilité que la réparation demandera plus que 2 heures.
b) Quelle est la probabilité que la réparation demandera plus que 5 heures, sachant qu'elle a déjà pris plus que 3 heures?

Réponses : (a) e^{-4}, (b) e^{-4}



Courbe de la densité de la loi normale centrée réduite

4.3. Loi normale

C'est la loi la plus importante de toutes les lois. Elle fut introduite par De Moivre en 1733 puis étudiée par Laplace et Gauss comme approximation de la loi binomiale pour n grand.

Définition 4.3. Une v.a. X suit une loi normale centrée réduite si sa densité est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \qquad \forall x \in \mathbb{R} \tag{11}$$

C'est en effet une densité de probabilité, puisque

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2}\, dt = 1$$

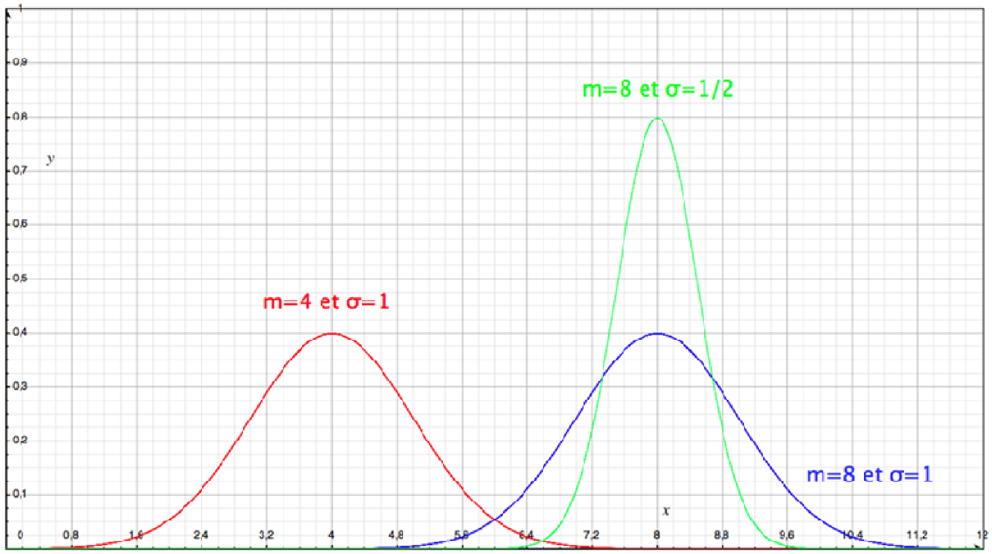
Proposition 4.1. *Soit X une v.a suivant une loi normale centrée réduite, alors*

$$\mathbb{E}(X) = 0 \qquad \mathrm{Var}(X) = 1 \tag{12}$$

Preuve. Il suffit d'intégrer par parties. ■

Définition 4.4. Une v.a. X suit une loi normale, ou de Laplace-Gauss, ou encore une loi gaussienne de paramètres m et σ, si sa densité est donnée par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \qquad \forall x \in \mathbb{R} \tag{13}$$



Courbe de la densité de la loi normale pour différentes valeurs des paramètres m et σ

On note $X \rightsquigarrow \mathcal{N}(m, \sigma)$.

Proposition 4.2. *Soit* $X \rightsquigarrow \mathcal{N}(m, \sigma)$, *alors*

$$\mathbb{E}(X) = m \qquad \mathrm{Var}(X) = \sigma^2 \tag{14}$$

Preuve. Procéder à un changement de variables $u = \frac{t-m}{\sigma}$ dans les intégrales et utiliser le résultat de la proposition 4.1. ■

Sur la figure ci-dessus, on voit bien que l'espérance indique la position de la gaussienne tandis que l'écart-type indique l'ampleur de sa dispersion

Une propriété importante de la famille des v.a. normales et qu'une transformation linéaire d'une v.a. normale donne toujours une v.a. normale.

Proposition 4.3. Si $X \rightsquigarrow \mathcal{N}(m, \sigma)$ alors $aX + b \rightsquigarrow \mathcal{N}(am + b, |a|\sigma)$, quelque soient $a \neq 0$ et b deux constantes réelles.

Preuve. Il suffit d'utiliser le résultat (*ii*) de la proposition 2.3. ■

Le corollaire important de ce résultat est que la v.a $X^* = \frac{X-m}{\sigma}$ suit une loi normale centrée réduite ($\mathcal{N}(0,1)$).

On ne peut pas exprimer en fonctions élémentaires la primitive de e^{-t^2}, et donc il est impossible de calculer la fonction de répartition. Par contre, ils existent des tables de valeurs de cette fonction (pour la loi $\mathcal{N}(0,1)$).

Notons Φ la fonction de répartition de la loi $\mathcal{N}(0,1)$. On a $\Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-t^2/2}\ dt$

Proposition 4.4. *Soit* Z *une v.a. normale centrée réduite. On a* $\Phi(-a) = 1 - \Phi(a)$. *En particulier* $\Phi(0) = 1/2$. *De plus*

$$\mathbb{P}(|Z| \leq a) = 2\Phi(a) - 1 \tag{15}$$

Preuve.

$$\begin{aligned}
\Phi(-a) + \Phi(a) &= \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{a} e^{-t^2/2}\ dt + \int_{-\infty}^{-a} e^{-t^2/2}\ dt\right) \\
&= \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{a} e^{-t^2/2}\ dt + \int_{a}^{+\infty} e^{-t^2/2}\ dt\right) \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-t^2/2}\ dt = 1
\end{aligned}$$

Où nous avons posé dans la deuxième intégrale $u = -t$. D'autre part

$$\mathbb{P}(|Z| \leq a) = \mathbb{P}(-a \leq Z \leq a) = \Phi(a) - \Phi(-a) = 2\Phi(a) - 1 \qquad \blacksquare$$

La table numérique de la loi centrée réduite (voir la dernière page de ce chapitre) donne les valeurs de $\Phi(a)$ pour a variant de 0 à 3.99. Cette table permet de calculer toute expression de la forme $\mathbb{P}(X \leq b)$,$\mathbb{P}(X \geq a)$ ou $\mathbb{P}(a \leq X \leq b)$. En effet

$$(X \leq b) = \left(\frac{X-m}{\sigma} \leq \frac{b-m}{\sigma}\right) \qquad \text{et} \qquad (X \geq a) = \left(\frac{X-m}{\sigma} \geq \frac{a-m}{\sigma}\right)$$

D'où

$$\mathbb{P}(X \leq b) = \mathbb{P}\left(X^* \leq \frac{b-m}{\sigma}\right) = \Phi\left(\frac{b-m}{\sigma}\right)$$

et

$$\mathbb{P}(X \geq a) = 1 - \Phi\left(\frac{a-m}{\sigma}\right)$$

Finalement

$$\mathbb{P}(a \leq X \leq b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$$

Utilisation de la table numérique. On lit les décimales dans les lignes, et les centièmes en colonnes. Par exemple, la valeur de $\Phi(1.65)$ se trouve à l'intersection de la ligne 1.6 et de la colonne 0.05. On trouve $\Phi(1.65) \cong 0.9505$, à 10^{-4} près. Pour les valeurs négatives de a, on utilise la relation $\Phi(-a) = 1 - \Phi(a)$.

Exemple. Soit $X \backsim \mathcal{N}(2,1)$. Calculer $\mathbb{P}(2 < X < 5)$.

La v.a $X^* = \frac{X-2}{1} = X - 2$ suit une loi normale centrée réduite. Donc

$$\mathbb{P}(2 < X < 5) = \mathbb{P}\left(\frac{2-2}{1} < X^* < \frac{5-2}{1}\right) = \mathbb{P}(0 < X < 3) = \Phi(3) - \Phi(0)$$
$$\cong 0.9986 - 0.5 \cong 0.4986$$

FIN DU CHAPITRE 6

Table de la loi normale centrée réduite

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8906	0.8925	0.8943	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9986	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000